09857826 .060701

JC18 Rec'd PCT/PTO 07 JUN 2001

FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		ATTORNEY'S DOCKET NUMBER **PF-0637 USN** U.S. APPLICATION NO. (If known, see 37 CFR 1.5) **TO BE ASSIGNED** 09/857826
INTERNATIONAL APPLICATION NO. **PCT/US99/30408**	INTERNATIONAL FILING DATE **10 December 1999**	PRIORITY DATE CLAIMED **11 December 1998**

TITLE OF INVENTION
NEURON-ASSOCIATED PROTEINS

APPLICANT(S) FOR DO/EO/US
INCYTE PHARMACEUTICALS, INC.; TANG, Y. Tom; YUE, Henry; BAUGHN, Mariah R.; HILLMAN, Jennifer L.; LAL, Preeti; AU-YOUNG, Janice; YANG, Junming; LU, Dyung Aina M.; AZIMZAI, Yalda

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is the **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)).
4. ☐ The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is attached hereto (required only if not communicated by the International Bureau)
   b. ☐ has been communicated by the International Bureau.
   c. ☒ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 16 below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.
13. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information:

1) Transmittal Letter (2 pp, in duplicate)
2) Return Postcard
3) Express Mail Label No.: **EL 856 153 993 US**
4) Request to Transfer

09857826 . 060701
531 Rec'd PCT 07 JUN 2001

U.S. APPLICATION NO. (if known, see 37 CFR 1.5) TO BE ASSIGNED 09/857826	INTERNATIONAL APPLICATION NO.: PCT/US99/30408	ATTORNEY'S DOCKET NUMBER PF-0637 USN

17. ☒ The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5):

Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO.......$1000.00

☐International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO..$860.00

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO........$710.00

☒**International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4)..........$690.00**

☐International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4)..................$100.00

ENTER APPROPRIATE BASIC FEE AMOUNT =				$690.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				$	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total Claims	20 =	0	X $ 18.00	$	
Independent Claims	2 =	0	X $ 80.00	$	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ $270.00	$	
TOTAL OF ABOVE CALCULATIONS =				$690.00	
☐ Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				$	
SUBTOTAL =				$690.00	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest clailmed priority date (37 CFR 1492(f)). +				$	
TOTAL NATIONAL FEE =				$690.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by the appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$	
TOTAL FEES ENCLOSED =				$690.00	
				Amount to be Refunded:	$
				Charged:	$

a. ☐ A check in the amount of $________ to cover the above fees is enclosed.
b. ☒ Please charge my Deposit Account No. 09-0108 in the amount of $ 690.00 to cover the above fees.
c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 09-0108. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

INCYTE GENOMICS, INC.
3160 Porter Drive
Palo Alto, CA 94304

D Hamlet-Cox
SIGNATURE

NAME: Diana Hamlet-Cox

REGISTRATION NUMBER: 33,302

DATE: 7 June 2001

PCT
WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/12, C07K 14/47, C12Q 1/68, A61K 38/17, C07K 16/18	A2	(11) International Publication Number:	WO 00/34477
		(43) International Publication Date:	15 June 2000 (15.06.00)

(21) International Application Number: PCT/US99/30408

(22) International Filing Date: 10 December 1999 (10.12.99)

(30) Priority Data:

Not furnished	11 December 1998 (11.12.98)	US
09/210,083	11 December 1998 (11.12.98)	US
60/119,365	9 February 1999 (09.02.99)	US
60/124,687	16 March 1999 (16.03.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	Not furnished (CIP)
Filed on	11 December 1998 (11.12.98)
US	60/119,365 (CIP)
Filed on	9 February 1999 (09.02.99)
US	60/124,687 (CIP)
Filed on	16 March 1999 (16.03.99)
US	09/210,083 (CIP)
Filed on	11 December 1998 (11.12.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). YANG, Junming [CN/US]; 7136 Clarendon Street, San Jose, CA 95129 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: NEURON–ASSOCIATED PROTEINS

```
1    M A - - - - - - - - - G S P S R A A G R R L Q L P - - - - -  2417014
1    M E F S L L L P R L E C N G A I S A H R N L R L P G S S D S  GI 3002527

17   - - - - - - - - - - - - - - - - - - - L L C L F L Q - - - -  2417014
31   P A S A S P V A G I T G M C T H A R L I L Y F F L V E M E F  GI 3002527

24   - - - G A T A V L F A V F - - - - - - - - - - V R Y N H K T  2417014
61   L H V G Q A G L E L P T S D D P S V S A S Q S A R Y R T G H  GI 3002527

41   D A A L - - - - - - - - - - - - - - - - - W H - - - - - - -  2417014
91   H A R L C L A N F C G R N R V S L M C P S W S P E L K Q S T  GI 3002527

47   - - - - - - - R S N H S N A D N E F Y F R Y - - P K E S H S  2417014
121  C L S L P K C W D Y R R A A V P G L F I L F F L R H R C P T  GI 3002527

68   V A Q A G V Q R R N L G S L Q P S P P R - - - - - - - - - -  2417014
151  L T Q D E V Q W C D H S S L Q P S T P E I K H P P A S A S Q  GI 3002527

88   - - - - - - - - - - - W - - - - - - - - - - - - S F A L V A  2417014
181  V A G T K D M H H Y T W L I F I F I F N F L R Q S L N S V T  GI 3002527
```

(57) Abstract

The invention provides human neuron–associated proteins (NEUAP) and polynucleotides which identify and encode NEUAP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of NEUAP.

NEURON-ASSOCIATED PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of neuron-associated proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative disorders including cancer; neuronal and neurological disorders; and autoimmune/inflammation disorders.

BACKGROUND OF THE INVENTION

The human nervous system, which regulates all bodily functions, is composed of the central nervous system (CNS), consisting of the brain and spinal cord, and the peripheral nervous system (PNS), consisting of afferent neural pathways for conducting nerve impulses from sensory organs to the CNS, and efferent neural pathways for conducting motor impulses from the CNS to effector organs. The PNS can be further divided into the somatic nervous system, which regulates voluntary motor activity such as for skeletal muscle, and the autonomic nervous system, which regulates involuntary motor activity for internal organs such as the heart, lungs, and viscera.

The central nervous system (CNS) is composed of more than 100 billion neurons at the spinal cord level, the lower brain level, and the higher brain or cortical level. Neurons transmit electric or chemical signals between cells. The spinal cord, a thin, tubular extension of the central nervous system within the bony spinal canal, contains ascending sensory and descending motor pathways, and is covered by membranes continuous with those of the brainstem and cerebral hemispheres. The spinal cord contains almost the entire motor output and sensory input systems of the trunk and limbs, and neuronal circuits in the cord also control rhythmic movements, such as walking, and a variety of reflexes. The lower areas of the brain such as the medulla, pons, mesencephalon, cerebellum, basal ganglia, substantia nigra, hypothalamus, and thalamus control unconscious activities including arterial pressure and respiration, equilibrium, and feeding reflexes, such as salivation. Emotions, such as anger, excitement, sexual response, and reaction to pain or pleasure, originate in the lower brain. The cerebral cortex or higher brain is the largest structure, consisting of a right and a left hemisphere interconnected by the corpus callosum. The cerebral cortex is involved in sensory, motor, and integrative functions related to perception, voluntary musculoskeletal movements, and the broad range of activities associated with consciousness, language, emotions, and memory. The cerebrum functions in association with the lower centers of the nervous system.

A nerve cell (neuron) contains four regions, the cell body, axon, dendrites, and axon terminal. The cell body contains the nucleus and other organelles. The dendrites are processes which extend

outward from the cell body and receive signals from sense organs or from the axons of other neurons. These signals are converted to electrical impulses and transmitted to the cell body. The axon, whose size can range from one millimeter to more than one meter, is a single process that conducts the nerve impulse away from the cell body. Cytoskeletal fibers, including microtubules and neurofilaments, run the length of the axon and function in transporting proteins, membrane vesicles, and other macromolecules from the cell body along the axon to the axon terminal. Some axons are surrounded by a myelin sheath made up of membranes from either an oligodendrocyte cell (CNS) or a Schwann cell (PNS). Myelinated axons conduct electrical impulses faster than unmyelinated ones of the same diameter. The axon terminal is at the tip of the axon away from the cell body. (See Lodish, H. et al. (1986) Molecular Cell Biology Scientific American Books New York NY, pp. 715-719.)

CNS-associated proteins have roles in neuronal signaling, cell adhesion, nerve regeneration, axon guidance, neurogenesis, and other functions. Certain CNS-associated proteins form an integral part of a membrane or are attached to a membrane. For example, neural membrane protein 35 (NMP35) is closely associated with neuronal membranes and is known to be highly expressed in the rat adult nervous system. (Schweitzer, B. et al. (1998) Mol. Cell. Neurosci. 11:260-273.) Synaptophysin (SY) is a major integral membrane protein of small synaptic vesicles. The chromosomal location of SY in human and mouse is on the X chromosome in subbands Xp11.22-p11.23. This region has been implicated in several inherited diseases including Wiskott-Aldrich syndrome, three forms of X-linked hypercalciuric nephrolithiaisis, and the eye disorders retinitis pigmentosa 2, congenital stationary night blindness, and Aland Island eye disease. (Fisher, S. E. et al. (1997) Genomics 45:340-347.) Peripherin or retinal degeneration slow protein (rds) is an integral membrane glycoprotein that is present in the rims of photoreceptor outer segment disks. In mammals, rds is thought to stabilize the disk rim through heterophilic interactions with related nonglycosylated proteins. Rds is a mouse neurological mutation that is characterized by abnormal development of rod and cone photoreceptors followed by their slow degeneration. (Kedzierski, W.J. et al. (1999) Neurochem. 72:430-438.)

43 KD postsynaptic protein or acteylcholine receptor-associated 43 KD protein (RAPSYN) is thought to play a role in anchoring or stabilizing the nicotinic acetylcholine receptor at synaptic sites. RAPSYN is involved in membrane association and may link the nicotinic acetylcholine receptor to the underlying postsynaptic cytoskeleton. (Buckel, A. et al. (1996) Genomics 35:613-616.) Neuritin is a protein whose gene is known to be induced by neural activity and by neurotrophins which promotes neuritogenesis. Neuraxin is a structural protein of the rat central nervous system that is believed to be immunologically related to microtubule-associated protein 5 (MAP5). Neuraxin is a novel type of neuron-specific protein which is characterized by an unusual amino acid composition,

12 central heptadecarepeats and putative protein and membrane interaction sites. The gene encoding neuraxin is unique in the haploid rat genome and is conserved in higher vertebrates. Neuraxin is implicated in neuronal membrane-microtubule interactions and is expressed throughout the rodent central nervous system (CNS). (Rienitz, A. et al. (1989) EMBO J. 8:2879-2888.)

NudC, a nuclear movement protein, interacts with the lissencephaly gene product Lis1, a protein involved in neuronal migration. People with Miller-Dieker syndrome (MDS) or isolated lissencephaly sequence (ILS) have a hemizygous deletion or mutation in the LIS1 gene. Both conditions are characterized by a smooth cerebral surface, a thickened cortex with four abnormal layers, and misplaced neurons. LIS1 is highly expressed in the ventricular zone and the cortical plate. The interaction of Lis1 with NudC, in conjunction with the MDS and ILS phenotypes, raises the possibility that nuclear movement in the ventricular zone is closely related to neuronal fates and to cortical architecture. (Morris, S. M. et al. (1998) Curr. Biol. 8:603-606.)

CNS-associated proteins can also be phosphoproteins. For example, ARPP-21 (cyclic AMP-regulated phosphoprotein) is a cytosolic neuronal phosphoprotein that is highly enriched in the striatum and in other dopaminoceptive regions of the brain. The steady-state level of ARPP-21 mRNA is developmentally regulated. But, in the neonatal and mature animal, ARPP-21 mRNA is not altered following 6-hydroxydopamine lesions of the substantia nigra or by pharmacologic treatments that upregulate the D1- or D2-dopamine receptors. (Ehrlich, M. E. et al. (1991) Neurochem. 57:1985-1991.)

CNS-associated signaling proteins may contain PDZ domains. PDZ domains have been found in proteins which act as adaptors in the assembly of multifunctional protein complexes involved in signaling events at surfaces of cell membranes. PDZ domains are generally found in membrane-associated proteins including neuronal nitric oxide synthase (NOS) and several dystrophin-associated proteins. (Ponting, C. P. et al. (1997) Bioessays 19:469-479.)

CNS-associated proteins may also contain epidermal growth factor (EGF) domains. The Notch proteins are transmembrane proteins which contain extracellular regions of repeated EGF domains. Notch proteins, such as the Drosophila melanogaster neurogenic protein Notch, are generally involved in the inhibition of developmental processes. Other members of the Notch family are the lin-12 and glp-1 genes of Caenorhabditis elegans. Genetic studies indicate that the lin-12 and glp-1 proteins act as receptors in specific developmental cell interactions which may be involved in certain embyronic defects. (Tax, F. E. et al. (1994) Nature 368:150-154.) Pecanex, a maternal-effect neurogenic locus of D. melanogaster is believed to encode a large transmembrane protein. In the absence of maternal expression of the pecanex gene, an embryo develops severe hyperneuralization similar to that characteristic of Notch mutant embryos. (LaBonne, S. G. et al. (1989) Dev. Biol.

136:1-116.) Other CNS-associated signaling proteins contain WW domains. The WW domain is a protein motif with two highly conserved tryptophans. It is present in a number of signaling and regulatory proteins, including Huntingtin interacting protein.

Alzheimer's disease (AD) is a degenerative disorder of the CNS which causes progressive memory loss and cognitive decline during mid to late adult life. AD is characterized by a wide range of neuropathologic features including amyloid deposits and intra-neuronal neurofibrillary tangles. Although the pathogenic pathway leading to neurodegeneration and AD is not well understood, at least three genetic loci that confer genetic susceptibility to the disease have been identified. (Schellenberg, G.D. (1995) Proc. Natl. Acad. Sci. 92:8552-8559; Sherrington, R. et al. (1995) Nature 375:754-760.)

Neuronal Thread Proteins (NTP) are a group of immunologically related molecules found in the brain and neuroectodermal tumor cell lines. NTP expression is increased in neuronal cells during proliferation, differentiation, brain development, in Alzheimer's disease (AD) brains, and in pathological states associated with regenerative nerve sprouting (de la Monte, S.M. et al. (1996) J. Neuropathol. Exp. Neurol. 55:1038-1050). Monoclonal antibodies generated to a recombinant NTP, AD7c-NTP, isolated from an end-stage AD brain library, showed high levels of NTP immunoreactivity in perikarya, neuropil fibers, and white matter fibers of AD brain tissue. In vitro studies also demonstrated NTP upregulation, phosphorylation, and translocation from the perikarya to cell processes and growth cones during growth factor-induced neuitic sprouting and neuronal differentiation. Additionally, increased NTP immunoreactivity was found in Down syndrome brains beginning in the second decade, prior to establishment of widespread AD neurodegeneration, and at an age when a low-level or an absence of NTP expression was observed in control brains. These findings indicated that abnormal expression and accumulation of NTP in brain may be an early marker of AD neurodegeneration in Down syndrome (de la Monte, S.M. et al. (1996) J. Neurol. Sci. 135:118-125). Furthermore, the increased expression and accumulation of NTP in AD brain tissue was paralleled by corresponding elevations of NTP in cerebrospinal fluid (CSF), and elevated levels of NTP were detectable in the CSF early in the course of the disease.

Astrocytomas, and the more malignant glioblastomas, are the most common primary tumors of the brain, accounting for over 65% of primary brain tumors. These tumors arise in glial cells of the astrocyte lineage. Following infection by pathogens, astrocytes function as antigen-presenting cells and modulate the activity of lymphocytes and macrophages. Astrocytomas constitutively express many cytokines and interleukins that are normally produced only after infection by a pathogen (de Micco, C. (1989) J. Neuroimmunol. 25:93-108). In the course of identifying genes related to astrocyte differentiation, one cDNA was isolated from an astrocytoma cDNA library that encodes a

protein structurally related to the plant pathogenesis-related (PR) proteins (Murphy, E.V. et al. (1995) Gene 159:131-135). The glioma pathogenesis-related protein (GliPR) is highly expressed in glioblastoma, but not in fetal or adult brain, or in other nervous system tumors. PR proteins are a family of small (10-20 kDa), protease resistant proteins induced in plants by viral infections, such as tobacco mosaic virus. The synthesis of PR proteins is believed to be part of a primitive immunological response in plants (van Loon, L.C. (1985) Plant Mol. Biol. 4:111-116). GliPR shares up to 50% homology with the PR-1 protein family over a region that comprises almost two thirds of the protein, including a conserved triad of amino acids, His-Glu-His, appropriately spaced to form a metal-binding domain (Murphy et al., supra).

Fe65-like protein (Fe65L2), a new member of the Fe65 protein family, is one of the ligands that interacts with the cytoplasmic domain of Alzheimer beta-amyloid precursor protein (APP). Transgenic mice expressing APP are known to simulate some of the prominent behavioral and pathological features of Alzheimer's disease, including age-related impairment in learning and memory, neuronal loss, gliosis, neuritic changes, amyloid deposition, and abnormal tau phosphorylation. Proteins that interact with the cytoplasmic domain of APP provide new insights into the physiological function of APP and, in turn, into the pathogenesis of Alzheimer's disease. (Duilio, A. et al. (1998) Biochem. J. 330:513-519.)

Contact from one neuron to another occurs at a specialized site called the synapse. At this site, the axon terminal from one neuron (the presynaptic cell) sends a signal to another neuron (the postsynaptic cell). Synapses may be connected either electrically or chemically. An electrical synapse consists of gap junctions connecting the two neurons, allowing electrical impulses to pass directly from the presynaptic to the postsynaptic cell. In a chemical synapse, the axon terminal of the presynaptic cell contains membrane vesicles containing a particular neurotransmitter molecule. A change in electrical potential at the nerve terminal resulting from the electrical impulse triggers the release of the neurotransmitter from the synaptic vesicle by exocytosis. The neurotransmitter rapidly diffuses across the synaptic cleft separating the presynaptic nerve cell from the postsynaptic cell. The neurotransmitter then binds receptors and opens transmitter-gated ion channels located in the plasma membrane of the postsynaptic cell, provoking a change in the cell's electrical potential. This change in membrane potential of the postsynaptic cell may serve either to excite or inhibit further transmission of the nerve impulse.

Neurotransmitters comprise a diverse group of some 30 small molecules which include acetylcholine, monoamines such as serotonin, dopamine, and histamine, and amino acids such as gamma-aminobutyric acid (GABA), glutamate, and aspartate, and neuropeptides such as endorphins and enkephalins. (McCance, K.L. and Huether, S.E. (1994) PATHOPHYSIOLOGY, The Biologic

Basis for Disease in Adults and Children, 2nd edition, Mosby, St. Louis, MO, pp 403-404.) Many of these molecules have more than one function and the effects may be excitatory, e.g. to depolarize the postsynaptic cell plasma membrane and stimulate nerve impulse transmission, or inhibitory, e.g. to hyperpolarize the plasma membrane and inhibit nerve impulse transmission.

Neurotransmitters and their receptors are targets of pharmacological agents aimed at controlling neurological function. For example GABA is the major inhibitory neurotransmitter in the CNS, and GABA receptors are the principal target of sedatives such as benzodiazepines and barbiturates which act by enhancing GABA-mediated effects (Katzung, B.G. (1995) Basic and Clinical Pharmacology, 6th edition, Appleton & Lange, Norwalk, CT, pp. 338-339). Diazepam binding inhibitor (DBI), also known as endozepine and acyl-Coenzyme (CoA)-binding protein, is an endogenous GABA receptor ligand which is thought to down-regulate the effects of GABA. DBI binds medium- and long-chain acyl-CoA esters with very high affinity and may function as an intracellular carrier of acyl-CoA esters (*125950 Diazepam Binding Inhibitor; DBI, Online Mendelian Inheritance in Man (OMIM); PROSITE PDOC00686 Acyl-CoA-binding protein signature). Aberrant activity of neurotransmitters and their receptors is involved in various neurological conditions, including Alzheimer's disease, myasthenia gravis, stroke, epilepsy, and Parkinson's disease. (See Planells-Cases, R. et al. (1993) Proc. Natl. Acad. Sci. USA 90:5057-5061.)

Each of over a trillion neurons in adult humans connects with over a thousand target cells (Tessier-Lavigne, M. et al. (1996) Science 274:1123-1133). These neuronal connections form during embryonic development. Each differentiating neuron sends out an axon tipped at the leading edge by a growth cone. Aided by molecular guidance cues, the growth cone migrates through the embryonic environment to its synaptic target. Semaphorins are growth cone guidance signals that may function during embryogenesis by providing local signals to specify territories inaccessible to growing axons (Puschel, A.W. et al. (1995) Neuron 14:941-948).

Axon growth is guided in part by contact-mediated mechanisms involving cell surface and extracellular matrix (ECM) molecules. Many ECM molecules, including fibronectin, vitronectin, members of the laminin, tenascin, collagen, and thrombospondin families, and a variety of proteoglycans, can act either as promoters or inhibitors of neurite outgrowth and extension (Tessier-Lavigne et al., supra). Receptors for ECM molecules include integrins, immunoglobulin superfamily members, and proteoglycans. ECM molecules and their receptors have also been implicated in the adhesion, maintenance, and differentiation of neurons (Reichardt, L.F. et al. (1991) Ann. Rev. Neurosci. 14:531-571). The proteoglycan testican is localized to the post-synaptic area of pyramidal cells of the hippocampus and may play roles in receptor activity, neuromodulation, synaptic plasticity, and neurotransmission (Bonnet, F. et al. (1996) J. Biol. Chem. 271:4373-4380).

Other nervous system-associated proteins have roles in neuron signaling, cell adhesion, nerve regeneration, axon guidance, and neurogenesis. The neurexophilins are neuropeptide-like proteins which are proteolytically processed after synthesis. They are ligands for the neuron-specific cell surface proteins, the α-neurexins. Neurexophilins and neurexins may participate in a neuron signaling pathway (Missler, M. and T.C. Sudhof (1998) J. Neurosci. 18:3630-3638; Missler, M. et al. (1998) J. Biol. Chem. 273:34716-34723). Ninjurin is a neuron cell surface protein which plays a role in cell adhesion and in nerve regeneration following injury. Ninjurin is up-regulated after nerve injury in dorsal root ganglion neurons and in Schwann cells (*602062 Ninjurin; NINJ1 OMIM; Araki, T. and Milbrandt, J. (1996) Neuron 17:353-361). Mammalian Numb is a phosphotyrosine-binding (PTB) domain-containing protein which may be involved in cortical neurogenesis and cell fate decisions in the mammalian nervous system. Numb's binding partner, the LNX protein, contains four PDZ domains and a ring finger domain and may participate in a signaling pathway involving Numb. PDZ domains have been found in proteins which act as adaptors in the assembly of multifunctional protein complexes involved in signaling events at surfaces of cell membranes (Ponting, C.P. (1997) Bioessays 19:469-479). LNX contains a tyrosine phosphorylation site which may be important for the binding of other PTB-containing proteins such as SHC, an adaptor protein which associates with tyrosine-phosphorylated growth factor receptors and downstream effectors (Dho, S.E. et al. (1998) J. Biol. Chem. 273:9179-9187).

The discovery of new neuron-associated proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative disorders including cancer; neuronal and neurological disorders; and autoimmune/inflammation disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, neuron-associated proteins, referred to collectively as "NEUAP" and individually as "NEUAP-1," "NEUAP-2," "NEUAP-3,""NEUAP-4,""NEUAP-5,""NEUAP-6,""NEUAP-7,""NEUAP-8,""NEUAP-9,""NEUAP-10,""NEUAP-11,""NEUAP-12,""NEUAP-13,""NEUAP-14,""NEUAP-15,""NEUAP-16,""NEUAP-17,""NEUAP-18,""NEUAP-19,""NEUAP-20," "NEUAP-21," "NEUAP-22," "NEUAP-23," "NEUAP-24," "NEUAP-25," "NEUAP-26," "NEUAP-27,", and "NEUAP-28." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-27 and fragments thereof. The invention also includes a polypeptide comprising an amino acid sequence that differs by one or more conservative amino acid substitutions from an amino acid sequence selected from the group consisting of SEQ ID NO:1-27.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-27 and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-27 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-27 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-27 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-27 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54 and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing a polynucleotide of the

invention under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-27 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-27 and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of NEUAP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-27 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of NEUAP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-27 and fragments thereof.

BRIEF DESCRIPTION OF THE FIGURES AND TABLES

Figures 1A and 1B show the amino acid sequence alignment between NEUAP-1 (2417014; SEQ ID NO:1) and a human neuronal thread protein, AD7c-NTP (GI 3002527; SEQ ID NO:55), produced using the multisequence alignment program of LASERGENE software (DNASTAR, Madison WI).

Figures 2A, 2B, and 2C show the amino acid sequence alignment between NEUAP-2 (2634931; SEQ ID NO:2) and a human glioma pathogenesis-related protein, GliPR (GI 847722; SEQ ID NO:56), produced using the multisequence alignment program of LASERGENE software.

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding NEUAP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of NEUAP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding NEUAP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze NEUAP, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"NEUAP" refers to the amino acid sequences of substantially purified NEUAP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of NEUAP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of NEUAP either by directly interacting with NEUAP or by acting on components of the biological pathway in which NEUAP participates.

An "allelic variant" is an alternative form of the gene encoding NEUAP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in

polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding NEUAP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as NEUAP or a polypeptide with at least one functional characteristic of NEUAP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding NEUAP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding NEUAP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent NEUAP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of NEUAP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of NEUAP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of NEUAP either by directly interacting with NEUAP or by acting on components of the biological pathway in which

NEUAP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind NEUAP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic NEUAP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid

(PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding NEUAP or fragments of NEUAP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of NEUAP or the polynucleotide encoding NEUAP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:28-54 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:28-54, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:28-54 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:28-54 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:28-54 and the region of SEQ ID NO:28-54 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-27 is encoded by a fragment of SEQ ID NO:28-54. A fragment of SEQ ID NO:1-27 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-27. For example, a fragment of SEQ ID NO:1-27 is useful as an immunogenic peptide

for the development of antibodies that specifically recognize SEQ ID NO:1-27. The precise length of a fragment of SEQ ID NO:1-27 and the region of SEQ ID NO:1-27 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms

is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62
Reward for match: 1
Penalty for mismatch: -2
Open Gap: 5 and Extension Gap: 2 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 11
Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a

standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into aecount conservative amino aeid substitutions. Sueh conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is seleeted as the default residue weight table. As with polynucleotide alignments, the pereent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequenee pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequenee, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequenees shown herein, in the tables, figures or Sequenee Listing, may be used to describe a length over whieh percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the proeess by which a polynucleotide strand anneals with a

complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 μg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid

support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of NEUAP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of NEUAP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding NEUAP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target

DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al.,1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for

example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding NEUAP, or fragments thereof, or NEUAP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers,

microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human neuron-associated proteins (NEUAP), the polynucleotides encoding NEUAP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative disorders including cancer; neuronal and neurological disorders; and autoimmune/inflammation disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding NEUAP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each NEUAP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each NEUAP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO: column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows the identity of each polypeptide; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

As shown in Figures 1A and 1B, NEUAP-1 has chemical and structural similarity with a human neuronal thread protein, AD7c-NTP (GI 3002527; SEQ ID NO:55). In particular, NEUAP-1 and AD7c-NTP share 24% identity, including a region of NEUAP-1 between residues S89 and Y127 in which the two proteins share 79% identity as well as two potential phosphorylation sites at S117 and S123.

MOTIFS, BLOCKS, and PFAM indicate that NEUAP-2 has an SCP-like extracellular protein signature, common to plant PR-1 proteins, between approximately residues S4 and G173. The conserved His-Glu-His triad of PR family proteins is found in NEUAP-2 at residues H78, E109, and H128. As shown in Figures 2A, 2B, and 2C, NEUAP-2 has chemical and structural similarity with a human glioma pathogenesis-related protein, GliPR (GI 847722; SEQ ID NO:56). In particular, the two proteins share 27% identity, the His-Glu-His triad, and ten of the twelve cysteine residues found in NEUAP-2, including C163, known to be involved in disulfide bond formation in PR-1 proteins.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions

associated with nucleotide sequences encoding NEUAP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:28-54 and to distinguish between SEQ ID NO:28-54 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express NEUAP as a fraction of total tissues expressing NEUAP. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing NEUAP as a fraction of total tissues expressing NEUAP. Of particular note is the expression of NDAP-2 in five neuronal tissues. Northern analysis shows the expression of NEUAP-1 in four tissues, three of which are cancerous, including a neuronal teratocarcinoma. Of particular interest is the tissue-specific expression of SEQ ID NO:31 and SEQ ID NO:32. SEQ ID NO:31 is highly expressed and SEQ ID NO:32 is exclusively expressed in nervous tissue. Of particular interest is the expression of SEQ ID NO:42, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO: 48, SEQ ID NO:52, SEQ ID NO:55, SEQ ID NO:56, and especially SEQ ID NO:51 in nervous tissues; and the expression of SEQ ID NO:42, SEQ ID NO:46, SEQ ID NO:48, and SEQ ID NO:51 in tissues associated with neurological disorders. Column 5 lists the vectors used to subclone each cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding NEUAP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses NEUAP variants. A preferred NEUAP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the NEUAP amino acid sequence, and which contains at least one functional or structural characteristic of NEUAP.

The invention also encompasses polynucleotides which encode NEUAP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:28-54, which encodes NEUAP.

The invention also encompasses a variant of a polynucleotide sequence encoding NEUAP. In particular, such a variant polynucleotide sequence will have at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding NEUAP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:28-54 which has at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting

of SEQ ID NO:28-54. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of NEUAP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding NEUAP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring NEUAP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode NEUAP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring NEUAP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding NEUAP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding NEUAP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode NEUAP and NEUAP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding NEUAP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:28-54 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-

Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding NEUAP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include

sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode NEUAP may be cloned in recombinant DNA molecules that direct expression of NEUAP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express NEUAP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter NEUAP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding NEUAP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, NEUAP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of NEUAP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid

chromatography. (See. e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins. Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active NEUAP, the nucleotide sequences encoding NEUAP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding NEUAP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding NEUAP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding NEUAP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding NEUAP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding NEUAP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding NEUAP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding NEUAP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding NEUAP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of NEUAP are needed, e.g. for the production of antibodies, vectors which direct high level expression of NEUAP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of NEUAP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of NEUAP. Transcription of sequences encoding NEUAP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding NEUAP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses NEUAP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma

virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of NEUAP in cell lines is preferred. For example, sequences encoding NEUAP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ and *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding NEUAP is inserted within a marker gene sequence, transformed cells containing

sequences encoding NEUAP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding NEUAP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding NEUAP and that express NEUAP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of NEUAP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on NEUAP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding NEUAP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding NEUAP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding NEUAP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence

and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode NEUAP may be designed to contain signal sequences which direct secretion of NEUAP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding NEUAP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric NEUAP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of NEUAP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the NEUAP encoding sequence and the heterologous protein sequence, so that NEUAP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled NEUAP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid

precursor, for example, ^{35}S-methionine.

Fragments of NEUAP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of NEUAP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of NEUAP and neuron-associated proteins. In addition, the expression of NEUAP is closely associated with nervous tissue, neurological disorders, cell proliferation including cancer, inflammation, and the immune response. Therefore, NEUAP appears to play a role in cell proliferative disorders including cancer; neuronal and neurological disorders; and autoimmune/inflammation disorders. In the treatment of disorders associated with increased NEUAP expression or activity, it is desirable to decrease the expression or activity of NEUAP. In the treatment of disorders associated with decreased NEUAP expression or activity, it is desirable to increase the expression or activity of NEUAP.

Therefore, in one embodiment, NEUAP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of NEUAP. Examples of such disorders include, but are not limited to, a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathesia, amnesia,

catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyenodocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma.

In another embodiment, a vector capable of expressing NEUAP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of NEUAP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified NEUAP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of NEUAP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of NEUAP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of NEUAP including, but not limited to, those listed above.

In a further embodiment, an antagonist of NEUAP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of NEUAP. Examples of such

disorders include, but are not limited to, those cell proliferative disorders including cancer; neuronal and neurological disorders; and autoimmune/inflammation disorders described above. In one aspect, an antibody which specifically binds NEUAP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express NEUAP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding NEUAP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of NEUAP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of NEUAP may be produced using methods which are generally known in the art. In particular, purified NEUAP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind NEUAP. Antibodies to NEUAP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with NEUAP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to NEUAP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of NEUAP

amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to NEUAP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce NEUAP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for NEUAP may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between NEUAP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering NEUAP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for NEUAP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of NEUAP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple NEUAP epitopes, represents the average affinity, or avidity, of the antibodies for NEUAP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular NEUAP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the NEUAP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of NEUAP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of NEUAP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding NEUAP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding NEUAP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding NEUAP. Thus, complementary molecules or fragments may be used to modulate NEUAP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding NEUAP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used

to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding NEUAP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding NEUAP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding NEUAP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding NEUAP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding NEUAP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques

for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding NEUAP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of NEUAP, antibodies to NEUAP, and mimetics, agonists, antagonists, or inhibitors of NEUAP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily

injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of NEUAP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example NEUAP or fragments thereof, antibodies of NEUAP, and agonists, antagonists or inhibitors of NEUAP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions

which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind NEUAP may be used for the diagnosis of disorders characterized by expression of NEUAP, or in assays to monitor patients being treated with NEUAP or agonists, antagonists, or inhibitors of NEUAP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for NEUAP include methods which utilize the antibody and a label to detect NEUAP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring NEUAP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of NEUAP expression. Normal or standard values for NEUAP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to NEUAP under conditions suitable for complex formation. The amount of standard complex

formation may be quantitated by various methods, such as photometric means. Quantities of NEUAP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding NEUAP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of NEUAP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of NEUAP, and to monitor regulation of NEUAP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding NEUAP or closely related molecules may be used to identify nucleic acid sequences which encode NEUAP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding NEUAP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the NEUAP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:28-54 or from genomic sequences including promoters, enhancers, and introns of the NEUAP gene.

Means for producing specific hybridization probes for DNAs encoding NEUAP include the cloning of polynucleotide sequences encoding NEUAP or NEUAP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding NEUAP may be used for the diagnosis of disorders associated with expression of NEUAP. Examples of such disorders include, but are not limited to, a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other

demyelinating diseases. bacterial and viral meningitis. brain abscess. subdural empyema. epidural abscess. suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru. Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathesia. amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, bipolar disorder, dementia, depression. Down's syndrome, peripheral neuropathy, bipolar disorder, dementia, depression, Down's syndrome, peripheral neuropathy, and Tourette's disorder; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyenodocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma. The

polynucleotide sequences encoding NEUAP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered NEUAP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding NEUAP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding NEUAP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding NEUAP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of NEUAP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding NEUAP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding NEUAP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding NEUAP, or a fragment of a polynucleotide complementary to the polynucleotide encoding NEUAP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of NEUAP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding NEUAP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra,

pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding NEUAP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, NEUAP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between NEUAP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with NEUAP, or fragments thereof, and washed. Bound NEUAP is then detected by methods well known in the art. Purified NEUAP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding NEUAP specifically compete with a test compound for binding NEUAP. In this manner, antibodies can be used to detect the presence of any peptide which shares

one or more antigenic determinants with NEUAP.

In additional embodiments, the nucleotide sequences which encode NEUAP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, are hereby expressly incorporated by reference.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/124,687, U.S. Ser. No. 60/119,365, and U.S. Ser. No. [Attorney Docket No. PF-0637 US, filed December 11, 1998], are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA

libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene

families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:28-54. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding NEUAP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Extension of NEUAP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:28-54 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this

fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on

antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:28-54 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:28-54 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array

elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the NEUAP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring NEUAP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of NEUAP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the NEUAP-encoding transcript.

IX. Expression of NEUAP

Expression and purification of NEUAP is achieved using bacterial or virus-based expression systems. For expression of NEUAP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3).

Antibiotic resistant bacteria express NEUAP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of NEUAP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding NEUAP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, NEUAP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from NEUAP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified NEUAP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of NEUAP Activity

NEUAP may be detected by the immunoreactivity of tissues to monoclonal antibodies (MAb) raised against recombinant NEUAP. Mabs to recombinant NEUAP may be prepared by methods well known in the art, and used to detect the expression of NEUAP in tissues by western blot analysis. Western blot analysis is carried out as described by de la Monte et al. (1996) J. Neuropathol. Exp. Neurol. supra. Cytosolic protein extracts of tissues are prepared and electrophoresed in SDS-PAGE Laemmli gels, and immunoblotted using Mabs raised against NEUAP. Antibody binding is detected with horseradish peroxidase-conjugated secondary antibody (IgG), and enhanced chemiluminesence reagents (Amersham Corp. Arlington Heights, IL). The amount of MAb immunoreactivity measured is proportional to the activity of NEUAP in the tissue preparation.

Alternatively, NEUAP, or biologically active fragments thereof, are labeled with ^{125}I

Bolton-Hunter reagent (see, eg., Bolton et al. (1973) Biochem. J. 133:529). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled NEUAP, washed, and any wells with labeled NEUAP complex are assayed. Data obtained using different concentrations of NEUAP are used to calculate values for the number, affinity, and association of NEUAP with the candidate molecules.

XI. Functional Assays

NEUAP function is assessed by expressing the sequences encoding NEUAP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of NEUAP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding NEUAP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding NEUAP and other genes of interest can be analyzed by northern

analysis or microarray techniques.

XII. Production of NEUAP Specific Antibodies

NEUAP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the NEUAP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-NEUAP activity by, for example, binding the peptide or NEUAP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring NEUAP Using Specific Antibodies

Naturally occurring or recombinant NEUAP is substantially purified by immunoaffinity chromatography using antibodies specific for NEUAP. An immunoaffinity column is constructed by covalently coupling anti-NEUAP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing NEUAP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of NEUAP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/NEUAP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and NEUAP is collected.

XIV. Identification of Molecules Which Interact with NEUAP

NEUAP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled NEUAP, washed, and any wells with labeled NEUAP complex are assayed. Data obtained using different concentrations of NEUAP are used to calculate values for the number, affinity, and

association of NEUAP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	28	2417014	HNT3AZT01	2417014H1, 2417014F6, and 2417014T6 (HNT3AZT01)
2	29	2634931	BONTNOT01	1599164X11 (BLADNOT03), 1840865R6 (COLNNOT07), 2634931H1 (BONTNOT01), 3016948F6 (MUSCNOT07), SBDA02985F1, SBDA03153F1, and SBAA04561F1
3	30	110960	PITUNOT01	110960F1, 110960H1, and 110960X31 (PITUNOT01), 1413173F6 (BRAINOT12), 2708730F6 (PONSAZT01)
4	31	380721	HYPONOB01	380721H1 (HYPONOB01), 530184R1 (BRAINOT03), 4313795H1 (BRAFNOT01)
5	32	829443	PROSTUT04	829443H1 and 829443T6 (PROSTUT04), 1356856F1 (LUNGNOT09), 1561879F1 (SPLNNOT04), 2454553F6 (ENDANOT01), 5113377H1 (ENDITXT01), SBDA04734F1
6	33	1470058	PANCTUT02	620887R6 (PGANNOT01), 667364R6 (SCORNOT01), 1001616R1 (BRSTNOT03), 1382686T1 (BRAITUT08), 3440580H2 (PENCNOT06), 4900807H1 (OVARDIT01)
7	34	1554947	BLADTUT04	444399R1 (MPHGNOT03), 1554947H1 (BLADTUT04), 2552447H1 (LUNGTUT06), 2776779H1 (PANCNOT15), 3140190H1 (SMCCNOT02), 3327533H1 (HEAONOT04), 4737377H1 (THYMNOR02)
8	35	1690245	PROSTUT10	380737R6 (HYPONOB01), 459187R6 (KERANOT01), 882488R1 (THYRNOT02), 1421177F1 (KIDNNOT09), 1690245F6 and 1690245H1 (PROSTUT10), 2880352H1 (UTRSTUT05)
9	36	1878262	LEUKNOT03	1878262F6 and 1878262H1 (LEUKNOT03), 3705684F6 (PENCNOT07)

Table 1 (Cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
10	37	2253519	OVARTUT01	917470R1 (BRSTNOT04), 1285941H1 (COLNNOT16), 1452424H1 (PENITUT01), 2101405H1 (BRAITUT02), 2253519H1, 2253519R6, and 2253519X308F2 (OVARTUT01), 2849605H1 (BRSTTUT13), 2941769F6 (PROSNOT28), 4540901H1 (THYRTMT01), 4699204F6 (BRALNOT01)
11	38	2888437	LUNGFET04	550739H1 (BEPINOT01), 639134R6 and 1004296R1 (BRSTNOT03), 1456837H1 (COLNFET02), 1576159F6 (LNODNOT03), 1813822F6 (SKINBIT01), 1965103R6 (BRSTNOT04), 2888437F6 and 2888437H1 (LUNGFET04), 3041589F6 and 3041589T6 (BRSTNOT16), 3316465F6 (PROSBPT03), 3416354H1 (PTHYNOT04), 3987261F6 (UTRSTUT05), 4527360H1 (LYMBTXT01)
12	39	3201753	PENCNOT02	060572X51 (LUNGNOT01), 1417168H1 (BRAINOT12), 1514580F1 (PANCTUT01), 1601609F6 (BLADNOT03), 1853144T6 (LUNGFET03), 2551341H1 (LUNGTUT06), 2967827F6 (SCORNOT04), 3201753F6 and 3201753H1 (PENCNOT02), 3435884F6 (PENCNOT05)
13	40	3800639	SPLNNOT12	152838R6 (FIBRAGT02), 820077H1 (KERANOT02), 1482425F1 (CORPNOT02), 1686313T6 (PROSNOT15), 1855749F6 (PROSNOT18), 2212060F6 (SINTFET03), 2679094H1 (SINIUCT01), 2685279H1 (LUNGNOT23), 2751789R6 (THP1AZS08), 3287040H1 (HEAONOT05), 3575146H1 (BRONNOT01), 3598393H1 (FIBPNOT01), 3798890H1 and 3800639H1 (SPLNNOT12), 4521233H1 (HNT2TXT01), 4988152F6 (LIVRTUT10), 5377369H1 and 5379308H1 (BRAXNOT01)
14	41	533825	BRAINOT03	533825H1 (BRAINOT03), 1661317F6 (BRSTNOT09), 3271477F6 (BRAINOT20), 3532613H1 (KIDNNOT25), 4338159H1 (BRAUNOT02), SBEA00478F1, SBEA02751F1

Table 1 (Cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
15	42	1311833	COLNFET02	1311833F6 (COLNFET02), 1311833H1 (COLNFET02), 1311833T1 (COLNFET02), 1492314H1 (PROSNON01), 1742220H1 (HIPONON01), 2279875R6 (PROSNON01), 2279875T6 (PROSNON01)
16	43	1342819	COLNTUT03	231227F1 (SINTNOT02), 1319329F1 (BLADNOT04), 1342819H1 (COLNTUT03), 1381830F1 (BRAITUT08), 3244424F7 (BRAINOT19)
17	44	1871288	SKINBIT01	1871288F6 (SKINBIT01), 1871288H1 (SKINBIT01), 1891163F6 (BLADTUT07)
18	45	2587338	BRAITUT22	2587338F6 (BRAITUT22), 2587338H1 (BRAITUT22)
19	46	2821211	ADRETUT06	2666281T6 (ADRETUT06), 2821211H1 (ADRETUT06), 2821211T6 (ADRETUT06), 2821626H1 (ADRETUT06), 3973838F6 (ADRETUT06)
20	47	2824832	ADRETUT06	2137150F6 (ENDCNOT01), 2137150T6 (ENDCNOT01), 2824832H1 (ADRETUT06), SBLA01910F1, SBLA01493F1, SBLA02371F1, SBLA01241F1
21	48	3070147	UTRSNOR01	1399942F1 (BRAITUT08), 3070147F6 (UTRSNOR01), 3070147H1 (UTRSNOR01)
22	49	3271841	BRAINOT20	531341F1 (BRAINOT03), 531341R6 (BRAINOT03), 1368113R1 (SCORNON02), 3271841H1 (BRAINOT20), 4227380F6 (BRAMDIT01)
23	50	3537827	SEMVNOT04	1376729F1 (LUNGNOT10), 1472735R6 (LUNGTUT03), 1995972T6 (BRSTTUT03), 2913592H1 (KIDNTUT15), 3174642F6 (UTRSTUT04), 3537827H1 (SEMVNOT04), 4261946F6 (BSCNDIT02), SBRA05006D1, SBRA01069D1
24	51	3729267	SMCCNON03	925471R1 (BRAINOT04), 988166R6 (LVENNOT03), 1303573F1 (PLACNOT02), 2176845F6 (ENDCNOT03), 3729267H1 (SMCCNON03)

Table 1 (Cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
25	52	3768771	BRSTNOT24	550415R6 (BEPINOT01), 1700822F6 (BLADTUT05), 1732040F6 (BRSTTUT08), 2028053R6 (KERANOT02), 2579651F6 (KIDNTUT13), 2731787F6 (OVARTUT04), 3447610H1 (THYMNOT08), 3498679H1 (PROSTUT13), 3606095H1 (LUNGNOT30), 3685266F6 (HEAANOT01), 3768771H1 (BRSTNOT24)
26	53	4248993	BRADDIR01	4248993F6 (BRADDIR01), 4248993H1 (BRADDIR01)
27	54	5402418	BRAHNOT01	270323F1 (HNT2NOT01), 950513T1 (PANCNOT05), 2083217F6 (UTRSNOT08), 2744356F6 (BRSTTUT14), 5402418H1 (BRAHNOT01)

Table 2

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods and Databases
1	198	S117, S51, S123	N49	M1-A27: Signal Peptide	AD7c-NTP (g3002527)	Motifs BLAST SPScan HMMER
2	463	T29, T50, T156, S195, S202, S299, T137, S218, S227, T323, S394	N114, N403, N409	M1-A27: Signal peptide Potential intramolecular disulfide-bridging site cysteine residues: C5, C74, C120, C126, C142, C147, C163, C183, C186, C192, C197, C283 SCP-like extracellular protein signature: S4-G173 PR family protein signature: H78, E109, H128	GliPR Human glioma pathogenesis-related protein (g847722)	Motifs SPScan BLAST BLOCKS PFAM

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods and Databases
3	316	S283 S37 T42 S74 S92 T125 T216 S285 T313	N191	G184-T210, V226-L244, V106-V121 and Y289-N314 Transmembrane Regions; V264-E316 Syntaptophysin Signature; G224-M277 Peripherin Signature	Neural Membrane Protein 35 (NMP35) (g3426268)	BLAST HMM BLOCKS
4	89	S56 T17 S33 S76			Cyclic AMP-Regulated Phosphoprotein (ARPP-21) (g238781)	BLAST
5	273	S96 S273 T54 S136 T190 S205 S252 S258 T64 S142 S268		C107-C134 and C141-C176 EGF-Like Domain Signature; R130-D132 Cell Attachment Sequence; C152-C163 Aspartic Acid and Asparagine Hydroxylation Site; M1-G19 Signal Peptide	Multiple EGF Protein (MEGF6) (g3449294)	BLAST Motifs SPScan HMM PFAM

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods and Databases
6	263	S216 S47 S109 S125 T126 S216 S248 S29 T95 T240		R97-D99 Cell Attachment Sequence; M1-S29 Signal Peptide	Fe65L2 Protein (g2293387)	BLAST Motifs SPScan
7	165	S44 T56 S108 T111 S137			Brain Expressed (BRX) Protein (g2196874)	BLAST
8	424	T373 S131 T257 T275 S284 T303 T322 T360 T361 T421 S312 Y266	N129	L251-P280 WW/rsp5/WWP Domain Signature	Huntingtin Interacting Protein (g3319282)	BLAST Motifs PFAM
9	164	S71 T129 T133		M1-A34 Signal Peptide	Neuritin (g2062678)	BLAST SPScan HMM
10	796	T147 S285 S353 S442 T475 S476 S591 S767 T135 S319 S383 S442 S543 T738 S753 S775 T780 Y60 Y133	N590	S661-G664, S704-G707 and S706-G709 Glycosaminoglycan Attachment Site	Pecanex Protein (SW:P18490)	BLAST Motifs

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods and Databases
11	854	T249 S398 T757 S88 T159 S175 T265 S330 S340 S387 S398 T557 S582 T594 S614 T626 T677 S712 T800 S99 T377 S494 T507 T649 T668 S750 Y422 Y593	N48 N153 N369 N375 N492 N561 N697 N747 N798	S373-Y422 Neuraxin Signature; L358-A409 43 Kd Postsynaptic Protein Signature	CNS Expressed Protein	Motifs BLOCKS
12	856	S370 T475 S604 S69 S71 S73 S238 T253 S284 S296 T414 T475 S625 T705 T835 T20 S119 S263 T337 T341 S386 T390 S599 S633 S634 T690 Y118	N18 N199 N369 N389 N531 N568 N721	M1-P65, M136-N218, H269-E349, T415-D497, T511-I592, T648-G733 and S773-S856 PDZ Domain Signature; R819-D821 Cell Attachment Sequence; S144-G147 Glycosaminoglycan Attachment Site	Brain Expressed Multi-PDZ Protein (g2959979)	BLAST Motifs Pfam
13	361	S120 T122 T197 T296 S48 S223 T243	N189 N264		MNUDC Protein (g2654358)	BLAST Motifs

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods and Databases
14	632	S48 S486 T549 S90 S91 S100 T159 S291 S292 T406 S431 S474 S574 S104 S107 T119 S124 S178 S191 S356 T397 Y440 Y499	N108 N157 N289 N384	PDZ domains: S178-E262; H285-Q367; V411-V496; D542-W627 Signal Peptide: M1-T18	LNXp70 (g3041881)	BLAST PFAM BLOCKS_PFAM SPScan MOTIFS
15	391	T77 S185 S203 S238 S36 T42 T63 S171 T191 T205 S223 T302 T334 S181 S220 S233	N324	Glycosaminoglycan attachment site: S85-G88 Protein Repeat Neurofilament: E123-K148; S111-P136 Transmembrane region: Y368-L388	heavy neurofilament subunit (g1841430)	BLOCKS_PRODOM BLAST MOTIFS HMM
16	490	S474 S90 T99 S105 S150 S269 S317 S335 S361 T26 S49 S87 T134 S238 S247 T255 T275 T329 T388 T401 T415 T454 S469 T476	N61 N189 N204 N359	Signal Peptide: M1-T19	neurofilament protein (g463250)	BLAST SPScan MOTIFS

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods and Databases
17	252	S116 S118 T155 T196 S219 S33 T81 S147 S164	N153	Glycosaminoglycan attachment site: S65-G68 Signal Peptide: M1-S33 Transmembrane region: I9-L27	bipolar disorder-associated protein (g2271473)	BLAST MOTIFS SPScan HMM
18	142	S3 S32 S36 T29	N106	Transmembrane region: L66-L84	ninjurin (g1644368)	BLAST HMM
19	67	T34			CNS-expressed protein (g862343)	BLAST
20	455	S39 T43 S104 T109 S185 S189 S204 S224 S226 S349 S365 T131 S358 S414	N58 N307	Acyl-CoA binding protein (DBI) signature: H41-P129	membrane-associated diazepam binding inhibitor MA-DBI (g244503)	PFAM BLOCKS PRINTS MOTIFS BLAST
21	252	S189 S54 S93 T119 Y242	N62 N127 N137 N143	Signal Peptide: M1-G22	neurexophilin (g508574)	BLAST SPScan

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods and Databases
22	149	S64 S40 S86 S10	N74	Transmembrane region: L125-L145	BM88 antigen; neuron-specific membrane protein (g557673)	BLAST HMM
23	204	S204 S16 T146 S157 T114	N6 N176	Transmembrane region: L66-A85 T114-W136	P24; neuron-specific membrane protein (g1890141)	BLAST HMM
24	367	S52 S55 T148 S181 T265 S303 T143 Y207		Glycosaminoglycan attachment site: S231-G234 Leucine-rich repeat: Y42-V88; K89-K134; E135-P180; S181-I228	leucine rich neuronal protein (g3135309)	PFAM BLOCKS BLAST MOTIFS
25	681	T70 T111 S155 S175 T206 S247 S375 S417 S502 T624 S633 T645 S653 S657 S66 T237 T349 S393 T539 S587 S647	N9 N254 N369 N474	Transmembrane region: C566-Y582	semaphorin (g854328)	BLAST HMM

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods and Databases
26	137	T11 T77 S134		Microbodies C-terminal targeting signal: A135-F137	myelin-associated/oligodendrocyte basic protein	BLAST MOTIFS
27	117		N81		GEF-2;ganglioside expression factor-2 (g2104570)	BLAST MOTIFS

Table 3

Nucleotide SEQ ID NO:	Selected Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
28	110-154 434-478		Cancer (0.750)	pINCY
29	1083-1127 1407-1451		Cancer (0.590)	pINCY
30	71-145	Nervous (0.727) Urologic (0.068) Endocrine (0.045)	Cell Proliferation (0.455) Neurological (0.205) Inflammation (0.227)	PBLUESCRIPT
31	379-438	Nervous (1.000)	Cell Proliferation (0.222) Inflammation (0.222) Neurological (0.222)	PBLUESCRIPT
32	255-314	Reproductive (0.222) Cardiovascular (0.176) Nervous (0.157)	Cell Proliferation (0.620) Inflammation (0.315)	PSPORT1
33	898-972	Reproductive (0.319) Nervous (0.191) Developmental (0.106) Gastrointestinal (0.106)	Cell Proliferation (0.766) Inflammation (0.298)	pINCY
34	258-317	Gastrointestinal (0.179) Reproductive (0.179) Cardiovascular (0.143) Nervous (0.143) Hematopoietic/Immune (0.143)	Cell Proliferation (0.572) Inflammation (0.393)	pINCY

Table 3 (Cont.)

Nucleotide SEQ ID NO:	Selected Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
35	184-243	Nervous (0.238) Reproductive (0.222) Gastrointestinal (0.127)	Cell Proliferation (0.603) Inflammation (0.254)	pINCY
36	306-380	Reproductive (0.500) Cardiovascular (0.250) Hematopoietic/Immune (0.250)	Cell Proliferation (0.750) Inflammation (0.250)	pINCY
37	88-147 433-507	Reproductive (0.355) Nervous (0.226) Musculoskeletal (0.129)	Cell Proliferation (0.677) Inflammation (0.129)	PSPORT1
38	83-142 1244-1318	Reproductive (0.230) Nervous (0.162) Cardiovascular (0.135)	Cell Proliferation (0.581) Inflammation (0.257)	pINCY
39	29-88 758-832	Nervous (0.281) Reproductive (0.246)	Cell Proliferation (0.491) Inflammation (0.228)	pINCY
40	435-494	Nervous (0.308) Reproductive (0.215) Hematopoietic/Immune (0.108)	Cell Proliferation (0.531) Inflammation (0.315) Neurological (0.100)	PINCY
41	1420-1482	Nervous (0.538) Reproductive (0.231) Urologic (0.077)	Cell Proliferation (0.462) Inflammation (0.346) Neurological (0.115)	pSPORT1
42	866-910	Reproductive (0.278) Developmental (0.222) Hematopoietic/Immune (0.167)	Cell Proliferation (0.611) Inflammation (0.389)	pINCY

Table 3 (Cont.)

Nucleotide SEQ ID NO:	Selected Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
43	948-992	Reproductive (0.287) Nervous (0.166) Cardiovascular (0.127) Hematopoietic/Immune (0.127)	Cell Proliferation (0.650) Inflammation (0.299)	pINCY
44	218-262	Reproductive (0.510) Gastrointestinal (0.143) Cardiovascular (0.102) Nervous (0.102)	Cell Proliferation (0.714) Inflammation (0.306)	pINCY
45	389-496	Nervous (0.267) Cardiovascular (0.200) Reproductive (0.200)	Cell Proliferation (0.533) Inflammation (0.333) Neurological (0.133)	pINCY
46	272-316	Nervous (0.667) Endocrine (0.167) Gastrointestinal (0.167)	Cell Proliferation (0.333) Inflammation (0.333)	pINCY
47	802-894	Developmental (0.267) Urologic (0.200) Endocrine (0.133) Reproductive (0.133)	Cell Proliferation (0.867) Inflammation (0.267)	pINCY
48	219-263	Reproductive (0.571) Cardiovascular (0.143) Nervous (0.143) Urologic (0.143)	Cell Proliferation (0.571) Inflammation (0.286)	pINCY

Table 3 (Cont.)

Nucleotide SEQ ID NO:	Selected Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
49	280-369	Nervous (0.885) Cardiovascular (0.038) Developmental (0.038) Endocrine (0.038)	Cell Proliferation (0.308) Inflammation (0.346) Neurological (0.269)	pINCY
50	487-531	Nervous (0.338) Reproductive (0.294) Urologic (0.088)	Cell Proliferation (0.647) Inflammation (0.221)	pINCY
51	337-420	Reproductive (0.250) Gastrointestinal (0.156) Nervous (0.156)	Cell Proliferation (0.500) Inflammation (0.438)	pINCY
52	412-474 928-1017	Reproductive (0.255) Gastrointestinal (0.170) Nervous (0.170)	Cell Proliferation (0.617) Inflammation (0.383)	pINCY
53	109-150	Nervous (1.000)	Inflammation (1.000)	pINCY
54	198-242	Nervous (0.281) Reproductive (0.156) Gastrointestinal (0.138)	Cell Proliferation (0.463) Inflammation (0.394)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
29	HNT3AZT01	Library was prepared from hNT precursor cells (at 80% confluence) treated for three days with 0.35 micromolar 5-aza-2'-deoxycytidine (a demethylating agent) in order to induce transcription of silent genes.
30	BONTNOT01	The library was constructed from normal bone connective tissue (periosteum) obtained from a 20-year-old Caucasian male during a hindquarter amputation. Pathology indicated partially necrotic and cystic osteoblastic grade 3 osteosarcoma (post chemotherapy) in the right lower limb. Patient history included osteogenesis imperfecta, bone infection of the lower limb, pathologic closed fracture, and non-union of fracture. Family history included osteogenesis imperfecta and closed fracture and diabetes with hyperosmolarity.
31	PITUNOT01	Library was constructed using RNA (Clontech, #6584-2, Lot 35278) obtained from the pituitary glands removed from a pool of 18 male and female Caucasian donors, 16 to 70 years old, who died from trauma.
32	HYPONOB01	Library was constructed using RNA (Clontech, #6579-2, Lot 3X843) isolated from the hypothalamus tissues of 51 male and female Caucasian donors, 16 to 75 years old.
33	PROSTUT04	Library was constructed using RNA isolated from prostate tumor tissue removed from a 57-year-old Caucasian male during radical prostatectomy, removal of both testes, and excision of regional lymph nodes. Pathology indicated adenocarcinoma (Gleason grade 3+3). Patient history included a benign neoplasm of the large bowel and type I diabetes. Family history included a malignant neoplasm of the prostate and type I diabetes.

Table 4 (Cont.)

Nucleotide SEQ ID NO:	Library	Library Description
34	PANCTUT02	Library was constructed using RNA isolated from pancreatic tumor tissue removed from a 45-year-old Caucasian female during radical pancreaticoduodenectomy. Pathology indicated a grade 4 anaplastic carcinoma. Family history included benign hypertension, hyperlipidemia, and atherosclerotic coronary artery disease.
35	BLADTUT04	Library was constructed using RNA isolated from bladder tumor tissue removed from a 60-year-old Caucasian male during a radical cystectomy, prostatectomy, and vasectomy. Pathology indicated grade 3 transitional cell carcinoma in the left bladder wall. Family history included type I diabetes, a malignant neoplasm of the stomach, atherosclerotic coronary artery disease, and an acute myocardial infarction.
36	PROSTUT10	Library was constructed using RNA isolated from prostatic tumor tissue removed from a 66-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated an adenocarcinoma (Gleason grade 2+3), and sadenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Family history included prostate cancer and secondary bone cancer.
37	LEUKNOT03	Library was constructed using RNA isolated frcm white blood cells of a 27-year-old female with blood type A+.

Table 4 (Cont.)

38	OVARTUT01	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 43-year-old Caucasian female during removal of the fallopian tubes and ovaries. Pathology indicated grade 2 mucinous cystadenocarcinoma involving the entire left ovary. Patient history included mitral valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, stress reaction, cerebrovascular disease, breast cancer, and uterine cancer.
Nucleotide SEQ ID NO:	Library	Library Description
39	LUNGFET04	Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 17 weeks' gestation from anencephalus.
40	PENCNOT02	Library was constructed using RNA isolated from penis right corpus cavernosum tissue removed from a male.
41	SPLNNOT12	Library was constructed using RNA isolated from spleen tissue removed from a 65-year-old female. Pathology indicated the spleen was negative for metastasis, and the associated tumor tissue indicated well-differentiated neuroendocrine carcinoma (islet cell tumor) forming a dominant mass in the distal pancreas.
42	BRAINOT03	Library was constructed using RNA isolated from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right fronto-parietal part of the brain.
43	COLNFET02	Library was constructed using RNA isolated from the colon tissue of a Caucasian female fetus, who died at 20 weeks' gestation.

Table 4 (Cont.)

44	COLNTUT03	Library was constructed using RNA isolated from colon tumor tissue obtained from the sigmoid colon of a 62-year-old Caucasian male during a sigmoidectomy and permanent colostomy. Pathology indicated invasive grade 2 adenocarcinoma. One lymph node contained metastasis with extranodal extension. Patient history included hyperlipidemia, cataract disorder, and dermatitis. Family history included benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, breast cancer, and prostate cancer.
45	SKINBIT01	Library was constructed using RNA isolated from diseased skin tissue of the left lower leg. Patient history included erythema nodosum of the left lower leg.
Nucleotide SEQ ID NO:	Library	Library Description
46	BRAITUT22	Library was constructed using RNA isolated from brain tumor tissue removed from the right frontal/parietal lobe of a 76-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a meningioma. Family history included senile dementia.
47	ADRETUT06	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 57-year-old Caucasian female during a unilateral right adrenalectomy. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.
48	ADRETUT06	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 57-year-old Caucasian female during a unilateral right adrenalectomy. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.
49	ADRETUT06	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 57-year-old Caucasian female during a unilateral right adrenalectomy. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.

Table 4 (Cont.)

50	UTRSNOR01	Library was constructed using RNA isolated from uterine endometrium tissue removed from a 29-year-old Caucasian female during a vaginal hysterectomy and cystocele repair. Pathology indicated the endometrium was secretory, and the cervix showed mild chronic cervicitis with focal squamous metaplasia. Pathology for the associated tumor tissue indicated intramural uterine leiomyoma. Patient history included hypothyroidism, pelvic floor relaxation, and paraplegia. Family history included benign hypertension, type II diabetes, and hyperlipidemia.

Table 4 (Cont.)

Nucleotide SEQ ID NO:	Library	Library Description
51	BRAINOT20	Library was constructed using RNA isolated from diseased brain tissue removed from the left temporal lobe of a 27-year-old Caucasian male during a brain lobectomy. Pathology for the left temporal lobe, including the mesial temporal structures, indicated focal, marked pyramidal cell loss and gliosis in hippocampal sector CA1, consistent with mesial temporal sclerosis. The left frontal lobe showed a focal deep white matter lesion, characterized by marked gliosis, calcifications, and hemosiderin-laden macrophages, consistent with a remote perinatal injury. This frontal lobe tissue also showed mild to moderate generalized gliosis, predominantly subpial and subcortical, consistent with chronic seizure disorder. GFAP was positive for astrocytes. Family history included brain cancer.
52	SEMVNOT04	Library was constructed using RNA isolated from seminal vesicle tissue removed from a 61-year-old Caucasian male during a radical prostatectomy. Pathology for the associated tumor tissue indicated adenocarcinoma, Gleason grade 3+3. The patient presented with induration, hyperplasia of the prostate, and elevated prostate specific antigen. Patient history included renal failure, osteoarthritis, left renal artery stenosis, thrombocytopenia, hyperlipidemia,and hepatitis C (carrier). Family history included benign hypertension.
53	SMCCNON03	This normalized smooth muscle cell library was constructed from 7.56 x 10^6 independent clones from a smooth muscle tissue library. Starting RNA was made from smooth muscle cell tissue removed from the coronary artery of a 3-year-old Caucasian male. The normalization and hybridization conditions were adapted from Soares et al., (Proc. Natl. Acad. Sci. USA (1994) 91:9228-9232); Swaroop et al., (Nucleic Acids Research (1991) 19:1954-806); and Bonaldo et al., (Genome Research (1996) 6:791-806), using a significantly longer (48 hour) reannealing hybridization period.

Table 4 (Cont.)

Nucleotide SEQ ID NO:	Library	Library Description
54	BRSTNOT24	Library was constructed using RNA isolated from diseased breast tissue removed from a 46-year-old Caucasian female during bilateral subcutaneous mammectomy. Pathology indicated nonproliferative fibrocystic disease bilaterally. Family history included breast cancer, benign hypertension, and atherosclerotic coronary artery disease.
55	BRADDIR01	Library was constructed using RNA isolated from diseased choroid plexus tissue of the lateral ventricle, removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema.
56	BRAHNOT01	Library was constructed using RNA isolated from posterior hippocampus tissue removed from a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Microscopically, the cerebral hemisphere revealed moderate fibrosis of the leptomeninges with focal calcifications. There was evidence of shrunken and slightly eosinophilic pyramidal neurons throughout the cerebral hemispheres. In addition, scattered throughout the cerebral cortex, there were multiple small microscopic areas of cavitation with surrounding gliosis. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly and an enlarged spleen and liver.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs*: Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs*: fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences*: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-27 and fragments thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54 and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of NEUAP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of NEUAP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/47	A2	(11) International Publication Number: WO 99/06554 (43) International Publication Date: 11 February 1999 (11.02.99)

(21) International Application Number: PCT/IB98/01238

(22) International Filing Date: 31 July 1998 (31.07.98)

(30) Priority Data:
08/905,134 1 August 1997 (01.08.97) US

(71) Applicant (*for all designated States except US*): GENSET [FR/FR]; 24, rue Royale, F-75008 Paris (FR).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): DUMAS MILNE EDWARDS, Jean-Baptiste [FR/FR]; 8, rue Grégoire de Tours, F-75006 Paris (FR). DUCLERT, Aymeric [FR/FR]; 6 ter, rue Victorine, F-94100 Saint-Maur (FR). LACROIX, Bruno [FR/FR]; 93, route de Vourles, F-69230 Saint-Genis Laval (FR).

(74) Agents: MARTIN, Jean-Jacques et al.; Cabinet Regimbeau, 26, avenue Kléber, F-75116 Paris (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: 5' ESTs FOR SECRETED PROTEINS EXPRESSED IN MUSCLE AND OTHER MESODERMAL TISSUES

(57) Abstract

The sequences of 5' ESTs derived from mRNAs encoding secreted proteins are disclosed. The 5' ESTs may be to obtain cDNAs and genomic DNAs corresponding to the 5' ESTs. The 5' ESTs may also be used in diagnostic, forensic, gene therapy, and chromosome mapping procedures. Upstream regulatory sequences may also be obtained using the 5' ESTs. The 5' ESTs may also be used to design expression vectors and secretion vectors.

```
1   M A - - - - - - - - - G S P S R A A G R R L Q L P - - - - -  2417014
1   M E F S L L L P R L E C N G A I S A H R N L R L P G S S D S  GI 3002527

17  - - - - - - - - - - - - - - - - - - - L L C L F L Q - - - -  2417014
31  P A S A S P V A G I T G M C T H A R L I L Y F F L V E M E F  GI 3002527

24  - - - G A T A V L F A V F - - - - - - - - - - V R Y N H K T  2417014
61  L H V G Q A G L E L P T S D D P S V S A S Q S A R Y R T G H  GI 3002527

41  D A A L - - - - - - - - - - - - - - - - - W H - - - - - - -  2417014
91  H A R L C L A N F C G R N R V S L M C P S W S P E L K Q S T  GI 3002527

47  - - - - - - - R S N H S N A D N E F Y F R Y - - P K E S H S  2417014
121 C L S L P K C W D Y R R A A V P G L F I L F F L R H R C P T  GI 3002527

68  V A Q A G V Q R R N L G S L Q P S P P R - - - - - - - - - -  2417014
151 L T Q D E V Q W C D H S S L Q P S T P E I K H P P A S A S Q  GI 3002527

88  - - - - - - - - - - - W - - - - - - - - - - - - S F A L V A  2417014
181 V A G T K D M H H Y T W L I F I F I F N F L R Q S L N S V T  GI 3002527
```

FIGURE 1A

```
95   QAGVQWHNLGSPQPLPPGFKRFSCLSLLSS 2417014
211  QAGVQWRNLGSLQPLPPGFKLFSCPSLLSS GI 3002527

125  WDYS----LESVFPLIAE------------ 2417014
241  WDYRRPPRLANFFVFLVEMGFTMFARLILI GI 3002527

139  ----------GQRSATSQAMHQL-----FG 2417014
271  SGPCDLPASASQSAGITGVSHHARLIFNFC GI 3002527

154  LFVTLMFASVGGG-----LGGL------LL 2417014
301  LFEMESHSVTQAGVQWPNLGSLQPLPPGLK GI 3002527

173  KLPFLDSPP-----RLPA--------LRGP 2417014
331  RFSCLSLPSSWDYGHLPPHPANFCIFIRGG GI 3002527

190  -SSLAGAWRA                     2417014
361  VSPYLSGWSQTPDLR                GI 3002527
```

FIGURE 1B

```
1   MHGSCSFLMLLLPLLLLLVATTGPVGALTD 2634931
1   M----------VSFVSNYSHTANILPDIEN GI 847722

31  EE-KRLMVELHNLYRAQVSPTASDMLHMRW 2634931
21  EDFIKDCVRIHNKFRSEVKPTASDMLYMTW GI 847722

60  DEELAAFAKAYARQCVWGHNKERGRRGENL 2634931
51  DPALAQIAKAWASNCQFSHNT-RLKPPHKL GI 847722

90  ---FAITDEGM---DVPL-----AMEEWHH 2634931
80  HPNFTSLGENIWTGSVPIFSVSSAITNWYD GI 847722

109 EREHYNLSAATCSPGQMCGHYTQVVWAKTE 2634931
110 EIQDYNFKTRICK--KVCGHYTQVVWADSY GI 847722

139 RIGCGSHFCEKLQGVEE-TNIELLVCNYEP 2634931
138 KVGCAVQFCPKVSGFDALSNGAHFICNYGP GI 847722

168 PGNVKGKRPYQEGTPCSQCPSGYHCKNSLC 2634931
168 GGNYPT-WPYKRGATCSACPNNDKCLDNLC GI 847722
```

FIGURE 2A

```
198 E P I G S P E D A Q D L P Y L V T E A P S F R A T E A S D S 2634931
197 - - - - - - - - - - - - - - - - - - - - - - - - V N D S E T GI 847722

228 R K M G T P S S L A T G I P A F L V T E V S G S L A T K A L 2634931
203 K - - - - - - - - - - - - - - - - - - - - - - - - - - - - - GI 847722

258 P A V E T Q A P T S L A T K D P P S M A T E A P P C V T T E 2634931
204 - - - - - - - - - - - - - - - - - - - - - - - - - - - - S N GI 847722

288 V P S I L A A H S L P S L D E E P V T F P K S T H V P I P K 2634931
206 V T T M L - - - - - - - - - - - - - - Y I R L A H I S T     GI 847722

318 S A D K V T D K T K V P S R S P E N S L D P K M S L T G A R 2634931
219                                                             GI 847722

348 E L L P H A Q E E A E A E A E L P P S S E V L A S V F P A Q 2634931
219                                                             GI 847722
```

FIGURE 2B

```
378 D K P G E L Q A T L D H T G H T S S K S L P N F P N T S A T 2634931
219                                                             GI 847722

408 A N A T G G R A L A L Q S S L P G A E G P D K P S V V S G L 2634931
219                                                             GI 847722

438 N S G P G H V W G P L L G L L L L P P L V L A G I F         2634931
219                                                             GI 847722
```

FIGURE 2C

OIPE JC184 OCT 1 5 2002 PATENT & TRADEMARK OFFICE

09857826 .060702

Docket No.: PF-0637 USN

Certificate of Mailing

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Box Sequence, Commissioner for Patents, P.O. Box 2327, Arlington, VA 22202 on October 9, 2002.

By: Joyce Abriam Printed: Joyce Abriam

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	Tang et al.		
Title:	NEURON-ASSOCIATED PROTEINS		
Serial No.:	09/857,826	Filing Date:	To Be Assigned
Examiner:	To Be Assigned	Group Art Unit:	To Be Assigned

Box Sequence
Commissioner for Patents
P.O. Box 2327
Arlington, VA 22202

CERTIFICATE UNDER 37 C.F.R. §3.73(b), REVOCATION OF POWER OF ATTORNEY AND APPOINTMENT OF NEW ATTORNEYS

Sir:

The undersigned has reviewed all the documents in the chain of title of the above-identified patent application and, to the best of undersigned's knowledge and belief, title is in the assignee identified above.

Incyte Genomics, Inc., formerly known as Incyte Pharmaceuticals, Inc., having a principal place of business located at 3160 Porter Drive, Palo Alto, California 94304, certifies that it is the assignee and owner of the entire right, title and interest in, to, and under the invention described and claimed in the above-identified application by virtue of an Assignment recorded at Reel 012605, Frame 0697, hereby revokes all previous powers of attorney and appoints the following patent attorneys/agents:

17

Lucy J. Billings	**Reg. No. 36,749**	**Shirley A. Recipon**	**Reg. No. 47,016**
Jenny Buchbinder	**Reg. No. 48,588**	**Cathleen M. Rocco**	**Reg. No. 46,172**
Michael C. Cerrone	**Reg. No. 39,132**	**Susan K. Sather**	**Reg. No. 44,316**
Diana Hamlet-Cox	**Reg. No. 33,302**	**Michelle M. Stempien**	**Reg. No. 41,327**
Joel Harris	**Reg. No. 44,743**	**David G. Streeter**	**Reg. No. 43,168**
Richard C. Ekstrom	**Reg. No. 37,027**	**Sreenivasarao Vepachedu**	**Reg. No. 46,395**
Barrie D. Greene	**Reg. No. 46,740**	**James M. Verna**	**Reg. No. 33,287**
Lori L. Kerber	**Reg. No. 41,113**	**Yu-Mei Eureka Wang**	**Reg. No. 50,510**
Lynn E. Murry	**Reg. No. 42,918**		

Docket No.: PF-0637 USN

Please direct all correspondence to:

Legal Department
Incyte Genomics, Inc.
3160 Porter Drive
Palo Alto, California 94304

and direct all telephone calls and facsimile transmissions to: Diana Hamlet-Cox, Incyte Genomics, Inc., Phone: (650) 845-4639, Fax: (650) 849-8886 or (650) 845-4166.

The undersigned (whose title is supplied below) is empowered to act on behalf of the assignee.

I hereby declare that all statements made herein of my own knowledge are true, and that these statements are made with the knowledge that willful false statements, and the like so made, are punishable by fine or imprisonment, or both, under Section 1001, Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

INCYTE GENOMICS, INC.

Date: October 9, 2002

By: Lee Bendekgey
Lee Bendekgey
EVP, General Counsel/Corporate Secretary

Docket No.: PF-0637 USN

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

NEURON-ASSOCIATED PROTEINS

the specification of which:

/____/ is attached hereto.

/ X __/ was filed on (to be assigned), as application Serial No. 09/857,826 and if this box contains an X /___/, was amended on ________________________________.

/ X __/ was filed as Patent Cooperation Treaty international application No. PCT/US99/30408 on December 10, 1999, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Article 19 on ______ 2001, and if this box contains an X /_/, was amended on ________________.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Country	Number	Filing Date	Priority Claimed
			/_/ Yes /_/ No
			/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/198,234	December 11, 1998	Expired
60/119,365	February 9, 1999	Expired
60/124,687	March 16, 1999	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)

I hereby appoint the following:

Lucy J. Billings	Reg. No. 36,749
Michael C. Cerrone	Reg. No. 39,132
Diana Hamlet-Cox	Reg. No. 33,302
Richard C. Ekstrom	Reg. No. 37,027
Barrie D. Greene	Reg. No. 46,740
Lynn E. Murry	Reg. No. 42,918
Shirley A. Recipon	Reg. No. 47,016
Susan K. Sather	Reg. No. 44,316
Michelle M. Stempien	Reg. No. 41,327
David G. Streeter	Reg. No. 43,168
Stephen Todd	Reg. No. 47,139
P. Ben Wang	Reg. No. 41,420

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

Docket No.: PF-0637 USN

LEGAL DEPARTMENT
INCYTE GENOMICS, INC.
3160 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

1-00

First Joint Inventor:

Full name: Y. Tom Tang

Signature: Y. Tom Tang

Date: Sept. 10, 2001

Citizenship: United States

Residence: San Jose, California CA

P.O. Address: 4230 Ranwick Court
San Jose, California 95118

Second Joint Inventor:

Full name: 2-00 Henry Yue

Signature: Henry Yue

Date: September 24, 2001

Citizenship: United States

Residence: Sunnyvale, California CA

P.O. Address: 826 Lois Avenue
Sunnyvale, California 94087

Docket No.: PF-0637 USN

Third Joint Inventor:

Full name: 300 Mariah R. Baughn

Signature: [illegible]

Date: 5 of September, 2001

Citizenship: United States

Residence: San Leandro, California CA

P.O. Address: 14244 Santiago Road
San Leandro, California 94577

Fourth Joint Inventor:

Full name: 400 Jennifer L. Hillman

Signature: [illegible]

Date: September 21, 2001

Citizenship: United States

Residence: Mountain View, California CA

P.O. Address: 230 Monroe Drive, #17
Mountain View, California 94040

Fifth Joint Inventor:

Full name: 500 Preeti Lal

Signature: Preeti Lal

Date: September 16, 2001

Citizenship: India

Residence: Santa Clara, California CA

P.O. Address: P.O. Box 5142
Santa Clara, California 95056

Docket No.: PF-0637 USN

Sixth Joint Inventor:

Full name: 6-00 Janice Au-Young

Signature:

Date: September 7, 2001

Citizenship: United States

Residence: Brisbane, California CA

P.O. Address: 233 Golden Eagle Lane
Brisbane, California 94005

Seventh Joint Inventor:

Full name: 7-00 Junming Yang

Signature:

Date: September 17, 2001

Citizenship: China

Residence: San Jose, California CA

P.O. Address: 7125 Bark Lane
San Jose, California 95129

Eighth Joint Inventor:

Full name: 8-00 Dyung Aina M. Lu

Signature:

Date: Sept 7, 2001

Citizenship: United States

Residence: San Jose, California CA

P.O. Address: 233 Coy Drive
San Jose, California 95123

9-00

Docket No.: PF-0637 USN

Ninth Joint Invendor:

Full name:	Yalda Azimzai
Signature:	Yalda Azimzai
Date:	September 13, 2001
Citizenship:	United States
Residence:	Castro Valley, California CA
P.O. Address:	5518 Boulder Canyon Drive Castro Valley, California 94552

SEQUENCE LISTING

```
<110> INCYTE PHARMACEUTICALS, INC.
      TANG, Y. Tom
      YUE, Henry
      BAUGHN, Mariah R.
      HILLMAN, Jennifer L.
      LAL, Preeti
      AU-YOUNG, Janice
      YANG, Junming
      LU, Dyung Aina M.
      AZIMZAI, Yalda

<120> NEURON-ASSOCIATED PROTEINS

<130> PF-0637 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/210,083; unassigned; 60/119,365; 60/124,687
<151> 1998-12-11; 1998-12-11; 1999-02-09; 1999-03-16

<160> 56

<170> PERL Program

<210> 1
<211> 198
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2417014CD1

<400> 1
Met Ala Gly Ser Pro Ser Arg Ala Ala Gly Arg Arg Leu Gln Leu
  1               5                  10                  15
Pro Leu Leu Cys Leu Phe Leu Gln Gly Ala Thr Ala Val Leu Phe
                 20                  25                  30
Ala Val Phe Val Arg Tyr Asn His Lys Thr Asp Ala Ala Leu Trp
                 35                  40                  45
His Arg Ser Asn His Ser Asn Ala Asp Asn Glu Phe Tyr Phe Arg
                 50                  55                  60
Tyr Pro Lys Glu Ser His Ser Val Ala Gln Ala Gly Val Gln Arg
                 65                  70                  75
Arg Asn Leu Gly Ser Leu Gln Pro Ser Pro Pro Arg Trp Ser Phe
                 80                  85                  90
Ala Leu Val Ala Gln Ala Gly Val Gln Trp His Asn Leu Gly Ser
                 95                 100                 105
Pro Gln Pro Leu Pro Pro Gly Phe Lys Arg Phe Ser Cys Leu Ser
                110                 115                 120
Leu Leu Ser Ser Trp Asp Tyr Ser Leu Glu Ser Val Phe Pro Leu
```

                125   130   135
Ile Ala Glu Gly Gln Arg Ser Ala Thr Ser Gln Ala Met His Gln
                140   145   150
Leu Phe Gly Leu Phe Val Thr Leu Met Phe Ala Ser Val Gly Gly
                155   160   165
Gly Leu Gly Gly Leu Leu Leu Lys Leu Pro Phe Leu Asp Ser Pro
                170   175   180
Pro Arg Leu Pro Ala Leu Arg Gly Pro Ser Ser Leu Ala Gly Ala
                185   190   195
Trp Arg Ala

<210> 2
<211> 463
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2634931CD1

<400> 2
Met His Gly Ser Cys Ser Phe Leu Met Leu Leu Leu Pro Leu Leu
 1   5   10   15
Leu Leu Leu Val Ala Thr Thr Gly Pro Val Gly Ala Leu Thr Asp
                20   25   30
Glu Glu Lys Arg Leu Met Val Glu Leu His Asn Leu Tyr Arg Ala
                35   40   45
Gln Val Ser Pro Thr Ala Ser Asp Met Leu His Met Arg Trp Asp
                50   55   60
Glu Glu Leu Ala Ala Phe Ala Lys Ala Tyr Ala Arg Gln Cys Val
                65   70   75
Trp Gly His Asn Lys Glu Arg Gly Arg Arg Gly Glu Asn Leu Phe
                80   85   90
Ala Ile Thr Asp Glu Gly Met Asp Val Pro Leu Ala Met Glu Glu
                95   100   105
Trp His His Glu Arg Glu His Tyr Asn Leu Ser Ala Ala Thr Cys
                110   115   120
Ser Pro Gly Gln Met Cys Gly His Tyr Thr Gln Val Val Trp Ala
                125   130   135
Lys Thr Glu Arg Ile Gly Cys Gly Ser His Phe Cys Glu Lys Leu
                140   145   150
Gln Gly Val Glu Glu Thr Asn Ile Glu Leu Leu Val Cys Asn Tyr
                155   160   165
Glu Pro Pro Gly Asn Val Lys Gly Lys Arg Pro Tyr Gln Glu Gly
                170   175   180
Thr Pro Cys Ser Gln Cys Pro Ser Gly Tyr His Cys Lys Asn Ser
                185   190   195
Leu Cys Glu Pro Ile Gly Ser Pro Glu Asp Ala Gln Asp Leu Pro
                200   205   210
Tyr Leu Val Thr Glu Ala Pro Ser Phe Arg Ala Thr Glu Ala Ser
                215   220   225
Asp Ser Arg Lys Met Gly Thr Pro Ser Ser Leu Ala Thr Gly Ile
                230   235   240
Pro Ala Phe Leu Val Thr Glu Val Ser Gly Ser Leu Ala Thr Lys

```
                    245                     250                     255
Ala Leu Pro Ala Val Glu Thr Gln Ala Pro Thr Ser Leu Ala Thr
                260                     265                     270
Lys Asp Pro Pro Ser Met Ala Thr Glu Ala Pro Pro Cys Val Thr
                275                     280                     285
Thr Glu Val Pro Ser Ile Leu Ala Ala His Ser Leu Pro Ser Leu
                290                     295                     300
Asp Glu Glu Pro Val Thr Phe Pro Lys Ser Thr His Val Pro Ile
                305                     310                     315
Pro Lys Ser Ala Asp Lys Val Thr Asp Lys Thr Lys Val Pro Ser
                320                     325                     330
Arg Ser Pro Glu Asn Ser Leu Asp Pro Lys Met Ser Leu Thr Gly
                335                     340                     345
Ala Arg Glu Leu Leu Pro His Ala Gln Glu Glu Ala Glu Ala Glu
                350                     355                     360
Ala Glu Leu Pro Pro Ser Ser Glu Val Leu Ala Ser Val Phe Pro
                365                     370                     375
Ala Gln Asp Lys Pro Gly Glu Leu Gln Ala Thr Leu Asp His Thr
                380                     385                     390
Gly His Thr Ser Ser Lys Ser Leu Pro Asn Phe Pro Asn Thr Ser
                395                     400                     405
Ala Thr Ala Asn Ala Thr Gly Gly Arg Ala Leu Ala Leu Gln Ser
                410                     415                     420
Ser Leu Pro Gly Ala Glu Gly Pro Asp Lys Pro Ser Val Val Ser
                425                     430                     435
Gly Leu Asn Ser Gly Pro Gly His Val Trp Gly Pro Leu Leu Gly
                440                     445                     450
Leu Leu Leu Leu Pro Pro Leu Val Leu Ala Gly Ile Phe
                455                     460


<210> 3
<211> 316
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 110960CD1

<400> 3
Met Thr Gln Gly Lys Leu Ser Val Ala Asn Lys Ala Pro Gly Thr
  1             5                   10                  15
Glu Gly Gln Gln Gln Val His Gly Glu Lys Lys Glu Ala Pro Ala
                20                  25                  30
Val Pro Ser Ala Pro Pro Ser Tyr Glu Glu Ala Thr Ser Gly Glu
                35                  40                  45
Gly Met Lys Ala Gly Ala Phe Pro Pro Ala Pro Thr Ala Val Pro
                50                  55                  60
Leu His Pro Ser Trp Ala Tyr Val Asp Pro Ser Ser Ser Ser Ser
                65                  70                  75
Tyr Asp Asn Gly Phe Pro Thr Gly Asp His Glu Leu Phe Thr Thr
                80                  85                  90
Phe Ser Trp Asp Asp Gln Lys Val Arg Arg Val Phe Val Arg Lys
```

```
                95                  100                 105
Val Tyr Thr Ile Leu Leu Ile Gln Leu Leu Val Thr Leu Ala Val
                110                 115                 120
Val Ala Leu Phe Thr Phe Cys Asp Pro Val Lys Asp Tyr Val Gln
                125                 130                 135
Ala Asn Pro Gly Trp Tyr Trp Ala Ser Tyr Ala Val Phe Phe Ala
                140                 145                 150
Thr Tyr Leu Thr Leu Ala Cys Cys Ser Gly Pro Arg Arg His Phe
                155                 160                 165
Pro Trp Asn Leu Ile Leu Leu Thr Val Phe Thr Leu Ser Met Ala
                170                 175                 180
Tyr Leu Thr Gly Met Leu Ser Ser Tyr Tyr Asn Thr Thr Ser Val
                185                 190                 195
Leu Leu Cys Leu Gly Ile Thr Ala Leu Val Cys Leu Ser Val Thr
                200                 205                 210
Val Phe Ser Phe Gln Thr Lys Phe Asp Phe Thr Ser Cys Gln Gly
                215                 220                 225
Val Leu Phe Val Leu Leu Met Thr Leu Phe Phe Ser Gly Leu Ile
                230                 235                 240
Leu Ala Ile Leu Leu Pro Phe Gln Tyr Val Pro Trp Leu His Ala
                245                 250                 255
Val Tyr Ala Ala Leu Gly Ala Gly Val Phe Thr Leu Phe Leu Ala
                260                 265                 270
Leu Asp Thr Gln Leu Leu Met Gly Asn Arg Arg His Ser Leu Ser
                275                 280                 285
Pro Glu Glu Tyr Ile Phe Gly Ala Leu Asn Ile Tyr Leu Asp Ile
                290                 295                 300
Ile Tyr Ile Phe Thr Phe Phe Leu Gln Leu Phe Gly Thr Asn Arg
                305                 310                 315
Glu


<210> 4
<211> 89
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 380721CD1

<400> 4
Met Ser Glu Gln Gly Asp Leu Asn Gln Ala Ile Ala Glu Glu Gly
 1              5                   10                  15
Gly Thr Glu Gln Glu Thr Ala Thr Pro Glu Asn Gly Ile Val Lys
                20                  25                  30
Ser Glu Ser Leu Asp Glu Glu Glu Lys Leu Glu Leu Gln Arg Arg
                35                  40                  45
Leu Glu Ala Gln Asn Gln Glu Arg Arg Lys Ser Lys Ser Gly Ala
                50                  55                  60
Gly Lys Gly Lys Leu Thr Arg Ser Leu Ala Val Cys Glu Glu Ser
                65                  70                  75
Ser Ala Arg Pro Gly Gly Glu Ser Leu Gln Gly Gln Thr Leu
                80                  85
```

```
<210> 5
<211> 273
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 829443CD1

<400> 5
Met Arg Gly Ser Gln Glu Val Leu Leu Met Trp Leu Leu Val Leu
  1               5                  10                  15
Ala Val Gly Gly Thr Glu His Ala Tyr Arg Pro Gly Arg Arg Val
                 20                  25                  30
Cys Ala Val Arg Ala His Gly Asp Pro Val Ser Glu Ser Phe Val
                 35                  40                  45
Gln Arg Val Tyr Gln Pro Phe Leu Thr Thr Cys Asp Gly His Arg
                 50                  55                  60
Ala Cys Ser Thr Tyr Arg Thr Ile Tyr Arg Thr Ala Tyr Arg Arg
                 65                  70                  75
Ser Pro Gly Leu Ala Pro Ala Arg Pro Arg Tyr Ala Cys Cys Pro
                 80                  85                  90
Gly Trp Lys Arg Thr Ser Gly Leu Pro Gly Ala Cys Gly Ala Ala
                 95                 100                 105
Ile Cys Gln Pro Pro Cys Arg Asn Gly Gly Ser Cys Val Gln Pro
                110                 115                 120
Gly Arg Cys Arg Cys Pro Ala Gly Trp Arg Gly Asp Thr Cys Gln
                125                 130                 135
Ser Asp Val Asp Glu Cys Ser Ala Arg Arg Gly Gly Cys Pro Gln
                140                 145                 150
Arg Cys Val Asn Thr Ala Gly Ser Tyr Trp Cys Gln Cys Trp Glu
                155                 160                 165
Gly His Ser Leu Ser Ala Asp Gly Thr Leu Cys Val Pro Lys Gly
                170                 175                 180
Gly Pro Pro Arg Val Ala Pro Asn Pro Thr Gly Val Asp Ser Ala
                185                 190                 195
Met Lys Glu Glu Val Gln Arg Leu Gln Ser Arg Val Asp Leu Leu
                200                 205                 210
Glu Glu Lys Leu Gln Leu Val Leu Ala Pro Leu His Ser Leu Ala
                215                 220                 225
Ser Gln Ala Leu Glu His Gly Leu Pro Asp Pro Gly Ser Leu Leu
                230                 235                 240
Val His Ser Phe Gln Gln Leu Gly Arg Ile Asp Ser Leu Ser Glu
                245                 250                 255
Gln Ile Ser Phe Leu Glu Glu Gln Leu Gly Ser Cys Ser Cys Lys
                260                 265                 270
Lys Asp Ser


<210> 6
<211> 263
<212> PRT
<213> Homo sapiens
```

<220>
<221> misc_feature
<223> Incyte ID No.: 1470058CD1

<400> 6
Met Leu Lys Cys His Val Phe Arg Cys Asp Val Pro Ala Lys Ala
1 5 10 15
Ile Ala Ser Ala Leu His Gly Leu Cys Ala Gln Ile Leu Ser Glu
20 25 30
Arg Val Glu Val Ser Gly Asp Ala Ser Cys Cys Ser Pro Asp Pro
35 40 45
Ile Ser Pro Glu Asp Leu Pro Arg Gln Val Glu Leu Leu Asp Ala
50 55 60
Val Ser Gln Ala Ala Gln Lys Tyr Glu Ala Leu Tyr Met Gly Thr
65 70 75
Leu Pro Val Thr Lys Ala Met Gly Met Asp Val Leu Asn Glu Ala
80 85 90
Ile Gly Thr Leu Thr Ala Arg Gly Asp Arg Asn Ala Trp Val Pro
95 100 105
Thr Met Leu Ser Val Ser Asp Ser Leu Met Thr Ala His Pro Ile
110 115 120
Gln Ala Glu Ala Ser Thr Glu Glu Glu Pro Leu Trp Gln Cys Pro
125 130 135
Val Arg Leu Val Thr Phe Ile Gly Val Gly Arg Asp Pro His Thr
140 145 150
Phe Gly Leu Ile Ala Asp Leu Gly Arg Gln Ser Phe Gln Cys Ala
155 160 165
Ala Phe Trp Cys Gln Pro His Ala Gly Gly Leu Ser Glu Ala Val
170 175 180
Gln Ala Ala Cys Met Val Gln Tyr Gln Lys Cys Leu Val Ala Ser
185 190 195
Ala Ala Arg Gly Lys Ala Trp Gly Ala Gln Ala Arg Ala Arg Leu
200 205 210
Arg Leu Lys Arg Thr Ser Ser Met Asp Ser Pro Gly Gly Pro Leu
215 220 225
Pro Leu Pro Leu Leu Lys Gly Gly Val Gly Gly Ala Gly Ala Thr
230 235 240
Pro Arg Lys Arg Gly Val Phe Ser Phe Leu Asp Ala Phe Arg Leu
245 250 255
Lys Pro Ser Leu Leu His Met Pro
260

<210> 7
<211> 165
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1554947CD1

<400> 7
Met Ala Asp Phe Asp Glu Ile Tyr Glu Glu Glu Glu Asp Glu Glu
1 5 10 15
Arg Ala Leu Glu Glu Gln Leu Leu Lys Tyr Ser Pro Asp Pro Val

20 25 30
Val Val Arg Gly Ser Gly His Val Thr Val Phe Gly Leu Ser Asn
35 40 45
Lys Phe Glu Ser Glu Phe Pro Ser Ser Leu Thr Gly Lys Val Ala
50 55 60
Pro Glu Glu Phe Lys Ala Ser Ile Asn Arg Val Asn Ser Cys Leu
65 70 75
Lys Lys Asn Leu Pro Val Asn Val Arg Trp Leu Leu Cys Gly Cys
80 85 90
Leu Cys Cys Cys Cys Thr Leu Gly Cys Ser Met Trp Pro Val Ile
95 100 105
Cys Leu Ser Lys Arg Thr Arg Arg Ser Ile Glu Lys Leu Leu Glu
110 115 120
Trp Glu Asn Asn Arg Leu Tyr His Lys Leu Cys Leu His Trp Arg
125 130 135
Leu Ser Lys Arg Lys Cys Glu Thr Asn Asn Met Met Glu Tyr Val
140 145 150
Ile Leu Ile Glu Phe Leu Pro Lys Thr Pro Ile Phe Arg Pro Asp
155 160 165

<210> 8
<211> 424
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1690245CD1

<400> 8
Met Gln Asn Leu Gly Met Thr Ser Pro Leu Pro Tyr Asp Ser Leu
1 5 10 15
Gly Tyr Asn Ala Pro His His Pro Phe Ala Gly Tyr Pro Pro Gly
20 25 30
Tyr Pro Met Gln Ala Tyr Val Asp Pro Ser Asn Pro Asn Ala Gly
35 40 45
Lys Val Leu Leu Pro Thr Pro Ser Met Asp Pro Val Cys Ser Pro
50 55 60
Ala Pro Tyr Asp His Ala Gln Pro Leu Val Gly His Ser Thr Glu
65 70 75
Pro Leu Ser Ala Pro Pro Pro Val Pro Val Val Pro His Val Ala
80 85 90
Ala Pro Val Glu Val Ser Ser Ser Gln Tyr Val Ala Gln Ser Asp
95 100 105
Gly Val Val His Gln Asp Ser Ser Val Ala Val Leu Pro Val Pro
110 115 120
Ala Pro Gly Pro Val Gln Gly Gln Asn Tyr Ser Val Trp Asp Ser
125 130 135
Asn Gln Gln Ser Val Ser Val Gln Gln Gln Tyr Ser Pro Ala Gln
140 145 150
Ser Gln Ala Thr Ile Tyr Tyr Gln Gly Gln Thr Cys Pro Thr Val
155 160 165
Tyr Gly Val Thr Ser Pro Tyr Ser Gln Thr Thr Pro Pro Ile Val
170 175 180

```
Gln Ser Tyr Ala Gln Pro Ser Leu Gln Tyr Ile Gln Gly Gln Gln
                185                 190                 195
Ile Phe Thr Ala His Pro Gln Gly Val Val Val Gln Pro Ala Ala
                200                 205                 210
Ala Val Thr Thr Ile Val Ala Pro Gly Gln Pro Gln Pro Leu Gln
                215                 220                 225
Pro Ser Glu Met Val Val Thr Asn Asn Leu Leu Asp Leu Pro Pro
                230                 235                 240
Pro Ser Pro Pro Lys Pro Lys Thr Ile Val Leu Pro Pro Asn Trp
                245                 250                 255
Lys Thr Ala Arg Asp Pro Glu Gly Lys Ile Tyr Tyr Tyr His Val
                260                 265                 270
Ile Thr Arg Gln Thr Gln Trp Asp Pro Pro Thr Trp Glu Ser Pro
                275                 280                 285
Gly Asp Asp Ala Ser Leu Glu His Glu Ala Glu Met Asp Leu Gly
                290                 295                 300
Thr Pro Thr Tyr Asp Glu Asn Pro Met Lys Ala Ser Lys Lys Pro
                305                 310                 315
Lys Thr Ala Glu Ala Asp Thr Ser Ser Glu Leu Ala Lys Lys Ser
                320                 325                 330
Lys Glu Val Phe Arg Lys Glu Met Ser Gln Phe Ile Val Gln Cys
                335                 340                 345
Leu Asn Pro Tyr Arg Lys Pro Asp Cys Lys Val Gly Arg Ile Thr
                350                 355                 360
Thr Thr Glu Asp Phe Lys His Leu Ala Arg Lys Leu Thr His Gly
                365                 370                 375
Val Met Asn Lys Glu Leu Lys Tyr Cys Lys Asn Pro Glu Asp Leu
                380                 385                 390
Glu Cys Asn Glu Asn Val Lys His Lys Thr Lys Glu Tyr Ile Lys
                395                 400                 405
Lys Tyr Met Gln Lys Phe Gly Ala Val Tyr Lys Pro Lys Glu Asp
                410                 415                 420
Thr Glu Leu Glu
```

<210> 9
<211> 164
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1878262CD1

<400> 9

```
Met Arg Cys Cys Arg Arg Arg Cys Cys Cys Arg Gln Pro Pro His
 1               5                  10                  15
Ala Leu Arg Pro Leu Leu Leu Leu Pro Leu Val Leu Leu Pro Pro
                20                  25                  30
Leu Ala Ala Ala Ala Ala Gly Pro Asn Arg Cys Asp Thr Ile Tyr
                35                  40                  45
Gln Gly Phe Ala Glu Cys Leu Ile Arg Leu Gly Asp Ser Met Gly
                50                  55                  60
Arg Gly Gly Glu Leu Glu Thr Ile Cys Arg Ser Trp Asn Asp Phe
```

```
                65                  70                  75
His Ala Cys Ala Ser Gln Val Leu Ser Gly Cys Pro Glu Glu Ala
                80                  85                  90
Ala Ala Val Trp Glu Ser Leu Gln Gln Glu Ala Arg Gln Ala Pro
                95                  100                 105
Arg Pro Asn Asn Leu His Thr Leu Cys Gly Ala Pro Val His Val
                110                 115                 120
Arg Glu Arg Gly Thr Gly Ser Lys Thr Asn Gln Glu Thr Leu Arg
                125                 130                 135
Ala Thr Ala Pro Ala Leu Pro Met Ala Pro Ala Pro Pro Leu Leu
                140                 145                 150
Ala Ala Ala Leu Ala Leu Ala Tyr Leu Leu Arg Pro Leu Ala
                155                 160


<210> 10
<211> 796
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2253519CD1

<400> 10
Met Thr Val Ala Gly Leu Lys Leu Leu Arg Ser Ala Phe Cys Cys
  1               5                  10                  15
Pro Pro Gln Gln Tyr Leu Thr Leu Ala Phe Thr Val Leu Leu Phe
                20                  25                  30
His Phe Asp Tyr Pro Arg Leu Ser Gln Gly Phe Leu Leu Asp Tyr
                35                  40                  45
Phe Leu Met Ser Leu Leu Cys Ser Lys Leu Trp Asp Leu Leu Tyr
                50                  55                  60
Lys Leu Arg Phe Val Leu Thr Tyr Ile Ala Pro Trp Gln Ile Thr
                65                  70                  75
Trp Gly Ser Ala Phe His Ala Phe Ala Gln Pro Phe Ala Val Pro
                80                  85                  90
His Ser Ala Met Leu Phe Val Gln Ala Leu Leu Ser Gly Leu Phe
                95                  100                 105
Ser Thr Pro Leu Asn Pro Leu Leu Gly Ser Ala Val Phe Ile Met
                110                 115                 120
Ser Tyr Ala Arg Pro Leu Lys Phe Trp Glu Arg Asp Tyr Asn Thr
                125                 130                 135
Lys Arg Val Asp His Ser Asn Thr Arg Leu Val Thr Gln Leu Asp
                140                 145                 150
Arg Asn Pro Gly Ala Asp Asp Asn Asn Leu Asn Ser Ile Phe Tyr
                155                 160                 165
Glu His Leu Thr Arg Ser Leu Gln His Thr Leu Cys Gly Asp Leu
                170                 175                 180
Val Leu Gly Arg Trp Gly Asn Tyr Gly Pro Gly Asp Cys Phe Val
                185                 190                 195
Leu Ala Ser Asp Tyr Leu Asn Ala Leu Val His Leu Ile Glu Val
                200                 205                 210
Gly Asn Gly Leu Val Thr Phe Gln Leu Arg Gly Leu Glu Phe Arg
                215                 220                 225
```

Gly Thr Tyr Cys Gln Gln Arg Glu Val Glu Ala Ile Thr Glu Gly
230 235 240
Val Glu Glu Asp Glu Gly Cys Cys Cys Cys Glu Pro Gly His Leu
245 250 255
Pro Arg Val Leu Ser Phe Asn Ala Ala Phe Gly Gln Arg Trp Leu
260 265 270
Ala Trp Glu Val Thr Ala Ser Lys Tyr Val Leu Glu Gly Tyr Ser
275 280 285
Ile Ser Asp Asn Asn Ala Ala Ser Met Leu Gln Val Phe Asp Leu
290 295 300
Arg Lys Ile Leu Ile Thr Tyr Tyr Val Lys Ser Ile Ile Tyr Tyr
305 310 315
Val Ser Arg Ser Pro Lys Leu Glu Val Trp Leu Ser His Glu Gly
320 325 330
Ile Thr Ala Ala Leu Arg Pro Val Arg Val Pro Gly Tyr Ala Asp
335 340 345
Ser Asp Pro Thr Phe Ser Leu Ser Val Asp Glu Asp Tyr Asp Leu
350 355 360
Arg Leu Ser Gly Leu Ser Leu Pro Ser Phe Cys Ala Val His Leu
365 370 375
Glu Trp Ile Gln Tyr Cys Ala Ser Arg Arg Thr Arg Pro Val Asp
380 385 390
Gln Asp Trp Asn Ser Pro Leu Val Thr Leu Cys Phe Gly Leu Cys
395 400 405
Val Leu Gly Arg Arg Ala Leu Gly Thr Ala Ser His Ser Met Ser
410 415 420
Ala Ser Leu Glu Pro Phe Leu Tyr Gly Leu His Ala Leu Phe Lys
425 430 435
Gly Asp Phe Arg Ile Thr Ser Pro Arg Asp Glu Trp Val Phe Ala
440 445 450
Asp Met Asp Leu Leu His Arg Val Val Ala Pro Gly Val Arg Met
455 460 465
Ala Leu Lys Leu His Gln Asp His Phe Thr Ser Pro Asp Glu Tyr
470 475 480
Glu Glu Pro Ala Ala Leu Tyr Asp Ala Ile Ala Ala Asn Glu Glu
485 490 495
Arg Leu Val Ile Ser His Glu Gly Asp Pro Ala Trp Arg Ser Ala
500 505 510
Ile Leu Ser Asn Thr Pro Ser Leu Leu Ala Leu Arg His Val Leu
515 520 525
Asp Asp Ala Ser Asp Glu Tyr Lys Ile Ile Met Leu Asn Arg Arg
530 535 540
His Leu Ser Phe Arg Val Ile Lys Val Asn Arg Glu Cys Val Arg
545 550 555
Gly Leu Trp Ala Gly Gln Gln Gln Glu Leu Val Phe Leu Arg Asn
560 565 570
Arg Asn Pro Glu Arg Gly Ser Ile Gln Asn Ala Lys Gln Ala Leu
575 580 585
Arg Asn Met Ile Asn Ser Ser Cys Asp Gln Pro Leu Gly Tyr Pro
590 595 600
Ile Tyr Val Ser Pro Leu Thr Thr Ser Leu Ala Gly Ser His Pro
605 610 615
Gln Leu Arg Ala Leu Trp Gly Gly Pro Ile Ser Leu Gly Ala Ile
620 625 630
Ala His Trp Leu Leu Arg Thr Trp Glu Arg Leu His Lys Gly Cys

```
                635                           640                           645
Gly Ala Gly Cys Asn Ser Gly Gly Asn Val Asp Asp Ser Asp Cys
                650                           655                           660
Ser Gly Gly Gly Gly Leu Thr Ser Leu Ser Asn Asn Pro Pro Val
                665                           670                           675
Ala His Pro Thr Pro Glu Asn Thr Ala Gly Asn Gly Asp Gln Pro
                680                           685                           690
Leu Pro Pro Gly Pro Gly Trp Gly Pro Arg Ser Ser Leu Ser Gly
                695                           700                           705
Ser Gly Asp Gly Arg Pro Pro Pro Leu Leu Gln Trp Pro Pro Pro
                710                           715                           720
Arg Leu Pro Gly Pro Pro Pro Ala Ser Pro Ile Pro Thr Glu Gly
                725                           730                           735
Pro Arg Thr Ser Arg Pro Pro Gly Pro Gly Leu Leu Ser Ser Glu
                740                           745                           750
Gly Pro Ser Gly Lys Trp Ser Leu Gly Gly Arg Lys Gly Leu Gly
                755                           760                           765
Gly Ser Asp Gly Glu Pro Ala Ser Gly Ser Pro Lys Gly Gly Thr
                770                           775                           780
Pro Lys Ser Gln Val Arg His Leu Trp Glu Gly Trp Val Pro Glu
                785                           790                           795
Gly


<210> 11
<211> 854
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2888437CD1

<400> 11
Met Lys Cys Leu Tyr Tyr Leu Tyr Ala Ser Leu Asp Pro Asn Ala
 1               5                  10                  15
Val Lys Ala Leu Asn Glu Met Trp Lys Cys Gln Asn Met Leu Arg
                20                  25                  30
Ile His Val Arg Glu Leu Leu Asp Leu His Lys Gln Pro Thr Ser
                35                  40                  45
Glu Ala Asn Cys Ser Ala Met Phe Gly Lys Leu Met Thr Ile Ala
                50                  55                  60
Lys Asn Leu Pro Asp Pro Gly Lys Ala Gln Asp Phe Val Lys Lys
                65                  70                  75
Phe Asn Gln Val Leu Gly Asp Asp Glu Lys Leu Arg Ser Gln Leu
                80                  85                  90
Glu Leu Leu Ile Ser Pro Thr Cys Ser Cys Lys Gln Ala Asp Ile
                95                 100                 105
Cys Val Arg Glu Ile Ala Arg Lys Leu Ala Asn Pro Lys Gln Pro
               110                 115                 120
Thr Asn Pro Phe Leu Glu Met Val Lys Phe Leu Leu Glu Arg Ile
               125                 130                 135
Ala Pro Val His Ile Asp Ser Glu Ala Ile Ser Ala Leu Val Lys
               140                 145                 150
```

```
Leu Met Asn Lys Ser Ile Glu Gly Thr Ala Asp Asp Glu Glu Glu
                155                 160                 165
Gly Val Ser Pro Asp Thr Ala Ile Arg Ser Gly Leu Glu Leu Leu
                170                 175                 180
Lys Val Leu Ser Phe Thr His Pro Thr Ser Phe His Ser Ala Glu
                185                 190                 195
Thr Tyr Glu Ser Leu Leu Gln Cys Leu Arg Met Glu Asp Asp Lys
                200                 205                 210
Val Ala Glu Ala Ala Ile Gln Ile Phe Arg Asn Thr Gly His Lys
                215                 220                 225
Ile Glu Thr Asp Leu Pro Gln Ile Arg Ser Thr Leu Ile Pro Ile
                230                 235                 240
Leu His Gln Lys Ala Lys Arg Gly Thr Pro His Gln Ala Lys Gln
                245                 250                 255
Ala Val His Cys Ile His Ala Ile Phe Thr Asn Lys Glu Val Gln
                260                 265                 270
Leu Ala Gln Ile Phe Glu Pro Leu Ser Arg Ser Leu Asn Ala Asp
                275                 280                 285
Val Pro Glu Gln Leu Ile Thr Pro Leu Val Ser Leu Gly His Ile
                290                 295                 300
Ser Met Leu Ala Pro Asp Gln Phe Ala Ser Pro Met Lys Ser Val
                305                 310                 315
Val Ala Asn Phe Ile Val Lys Asp Leu Leu Met Asn Asp Arg Ser
                320                 325                 330
Thr Gly Glu Lys Asn Gly Lys Leu Trp Ser Pro Asp Glu Glu Val
                335                 340                 345
Ser Pro Glu Val Leu Ala Lys Val Gln Ala Ile Lys Leu Leu Val
                350                 355                 360
Arg Trp Leu Leu Gly Met Lys Asn Asn Gln Ser Lys Ser Ala Asn
                365                 370                 375
Ser Thr Leu Arg Leu Leu Ser Ala Met Leu Val Ser Glu Gly Asp
                380                 385                 390
Leu Thr Glu Gln Lys Arg Ile Ser Lys Ser Asp Met Ser Arg Leu
                395                 400                 405
Arg Leu Ala Ala Gly Ser Ala Ile Met Lys Leu Ala Gln Glu Pro
                410                 415                 420
Cys Tyr His Glu Ile Ile Thr Pro Glu Gln Phe Gln Leu Cys Ala
                425                 430                 435
Leu Val Ile Asn Asp Glu Cys Tyr Gln Val Arg Gln Ile Phe Ala
                440                 445                 450
Gln Lys Leu His Lys Ala Leu Val Lys Leu Leu Leu Pro Leu Glu
                455                 460                 465
Tyr Met Ala Ile Phe Ala Leu Cys Ala Lys Asp Pro Val Lys Glu
                470                 475                 480
Arg Arg Ala His Ala Arg Gln Cys Leu Leu Lys Asn Ile Ser Ile
                485                 490                 495
Arg Arg Glu Tyr Ile Lys Gln Asn Pro Met Ala Thr Glu Lys Leu
                500                 505                 510
Leu Ser Leu Leu Pro Glu Tyr Val Val Pro Tyr Met Ile His Leu
                515                 520                 525
Leu Ala His Asp Pro Asp Phe Thr Arg Ser Gln Asp Val Asp Gln
                530                 535                 540
Leu Arg Asp Ile Lys Glu Cys Leu Trp Phe Met Leu Glu Val Leu
                545                 550                 555
Met Thr Lys Asn Glu Asn Asn Ser His Ala Phe Met Lys Lys Met
```

```
                    560                      565                      570
Ala Glu Asn Ile Lys Leu Thr Arg Asp Ala Gln Ser Pro Asp Glu
                    575                      580                      585
Ser Lys Thr Asn Glu Lys Leu Tyr Thr Val Cys Asp Val Ala Leu
                    590                      595                      600
Cys Val Ile Asn Ser Lys Ser Ala Leu Cys Asn Ala Asp Ser Pro
                    605                      610                      615
Lys Asp Pro Val Leu Pro Met Lys Phe Phe Thr Gln Pro Glu Lys
                    620                      625                      630
Asp Phe Cys Asn Asp Lys Ser Tyr Ile Ser Glu Glu Thr Arg Val
                    635                      640                      645
Leu Leu Leu Thr Gly Lys Pro Lys Pro Ala Gly Val Leu Gly Ala
                    650                      655                      660
Val Asn Lys Pro Leu Ser Ala Thr Gly Arg Lys Pro Tyr Val Arg
                    665                      670                      675
Ser Thr Gly Thr Glu Thr Gly Ser Asn Ile Asn Val Asn Ser Glu
                    680                      685                      690
Leu Asn Pro Ser Thr Gly Asn Arg Ser Arg Glu Gln Ser Ser Glu
                    695                      700                      705
Ala Ala Glu Thr Gly Val Ser Glu Asn Glu Glu Asn Pro Val Arg
                    710                      715                      720
Ile Ile Ser Val Thr Pro Val Lys Asn Ile Asp Pro Val Lys Asn
                    725                      730                      735
Lys Glu Ile Asn Ser Asp Gln Ala Thr Gln Gly Asn Ile Ser Ser
                    740                      745                      750
Asp Arg Gly Lys Lys Arg Thr Val Thr Ala Ala Gly Ala Glu Asn
                    755                      760                      765
Ile Gln Gln Lys Thr Asp Glu Lys Val Asp Glu Ser Gly Pro Pro
                    770                      775                      780
Ala Pro Ser Lys Pro Arg Arg Gly Arg Arg Pro Lys Ser Glu Ser
                    785                      790                      795
Gln Gly Asn Ala Thr Lys Asn Asp Asp Leu Asn Lys Pro Ile Asn
                    800                      805                      810
Lys Gly Arg Lys Arg Ala Ala Val Gly Gln Glu Ser Pro Gly Gly
                    815                      820                      825
Leu Glu Ala Gly Asn Ala Lys Ala Pro Lys Leu Gln Asp Leu Ala
                    830                      835                      840
Lys Lys Ala Ala Pro Ala Glu Arg Gln Ile Asp Leu Gln Arg
                    845                      850


<210> 12
<211> 856
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3201753CD1

<400> 12
Met Arg Gly Ile Phe Ile Lys His Val Leu Glu Asp Ser Pro Ala
  1               5                   10                  15
Gly Lys Asn Gly Thr Leu Lys Pro Gly Asp Arg Ile Val Glu Val
                 20                   25                  30
```

Asp Gly Met Asp Leu Arg Asp Ala Ser His Glu Gln Ala Val Glu
35 40 45
Ala Ile Arg Lys Ala Gly Asn Pro Val Val Phe Met Val Gln Ser
50 55 60
Ile Ile Asn Arg Pro Arg Ala Pro Ser Gln Ser Glu Ser Glu Pro
65 70 75
Glu Lys Ala Pro Leu Cys Ser Val Pro Pro Pro Pro Pro Ser Ala
80 85 90
Phe Ala Glu Met Gly Ser Asp His Thr Gln Ser Ser Ala Ser Lys
95 100 105
Ile Ser Gln Asp Val Asp Lys Glu Asp Glu Phe Gly Tyr Ser Trp
110 115 120
Lys Asn Ile Arg Glu Arg Tyr Gly Thr Leu Thr Gly Glu Leu His
125 130 135
Met Ile Glu Leu Glu Lys Gly His Ser Gly Leu Gly Leu Ser Leu
140 145 150
Ala Gly Asn Lys Asp Arg Ser Arg Met Ser Val Phe Ile Val Gly
155 160 165
Ile Asp Pro Asn Gly Ala Ala Gly Lys Asp Gly Arg Leu Gln Ile
170 175 180
Ala Asp Glu Leu Leu Glu Ile Asn Gly Gln Ile Leu Tyr Gly Arg
185 190 195
Ser His Gln Asn Ala Ser Ser Ile Ile Lys Cys Ala Pro Ser Lys
200 205 210
Val Lys Ile Ile Phe Ile Arg Asn Lys Asp Ala Val Asn Gln Met
215 220 225
Ala Val Cys Pro Gly Asn Ala Val Glu Pro Leu Pro Ser Asn Ser
230 235 240
Glu Asn Leu Gln Asn Lys Glu Thr Glu Pro Thr Val Thr Thr Ser
245 250 255
Asp Ala Ala Val Asp Leu Ser Ser Phe Lys Asn Val Gln His Leu
260 265 270
Glu Leu Pro Lys Asp Gln Gly Gly Leu Gly Ile Ala Ile Ser Glu
275 280 285
Glu Asp Thr Leu Ser Gly Val Ile Ile Lys Ser Leu Thr Glu His
290 295 300
Gly Val Ala Ala Thr Asp Gly Arg Leu Lys Val Gly Asp Gln Ile
305 310 315
Leu Ala Val Asp Asp Glu Ile Val Val Gly Tyr Pro Ile Glu Lys
320 325 330
Phe Ile Ser Leu Leu Lys Thr Ala Lys Met Thr Val Lys Leu Thr
335 340 345
Ile His Ala Glu Asn Pro Asp Ser Gln Ala Val Pro Ser Ala Ala
350 355 360
Gly Ala Ala Ser Gly Glu Lys Lys Asn Ser Ser Gln Ser Leu Met
365 370 375
Val Pro Gln Ser Gly Ser Pro Glu Pro Glu Ser Ile Arg Asn Thr
380 385 390
Ser Arg Ser Ser Thr Pro Ala Ile Phe Ala Ser Asp Pro Ala Thr
395 400 405
Cys Pro Ile Ile Pro Gly Cys Glu Thr Thr Ile Glu Ile Ser Lys
410 415 420
Gly Arg Thr Gly Leu Gly Leu Ser Ile Val Gly Gly Ser Asp Thr
425 430 435
Leu Leu Gly Ala Ile Ile Ile His Glu Val Tyr Glu Glu Gly Ala

440 445 450
Ala Cys Lys Asp Gly Arg Leu Trp Ala Gly Asp Gln Ile Leu Glu
455 460 465
Val Asn Gly Ile Asp Leu Arg Lys Ala Thr His Asp Glu Ala Ile
470 475 480
Asn Val Leu Arg Gln Thr Pro Gln Arg Val Arg Leu Thr Leu Tyr
485 490 495
Arg Asp Glu Ala Pro Tyr Lys Glu Glu Glu Val Cys Asp Thr Leu
500 505 510
Thr Ile Glu Leu Gln Lys Lys Pro Gly Lys Gly Leu Gly Leu Ser
515 520 525
Ile Val Gly Lys Arg Asn Asp Thr Gly Val Phe Val Ser Asp Ile
530 535 540
Val Lys Gly Gly Ile Ala Asp Ala Asp Gly Arg Leu Met Gln Gly
545 550 555
Asp Gln Ile Leu Met Val Asn Gly Glu Asp Val Arg Asn Ala Thr
560 565 570
Gln Glu Ala Val Ala Ala Leu Leu Lys Cys Ser Leu Gly Thr Val
575 580 585
Thr Leu Glu Val Gly Arg Ile Lys Ala Gly Pro Phe His Ser Glu
590 595 600
Arg Arg Pro Ser Gln Ser Ser Gln Val Ser Glu Gly Ser Leu Ser
605 610 615
Ser Phe Thr Phe Pro Leu Ser Gly Ser Ser Thr Ser Glu Ser Leu
620 625 630
Glu Ser Ser Ser Lys Lys Asn Ala Leu Ala Ser Glu Ile Gln Gly
635 640 645
Leu Arg Thr Val Glu Met Lys Lys Gly Pro Thr Asp Ser Leu Gly
650 655 660
Ile Ser Ile Ala Gly Gly Val Gly Ser Pro Leu Gly Asp Val Pro
665 670 675
Ile Phe Ile Ala Met Met His Pro Thr Gly Val Ala Ala Gln Thr
680 685 690
Gln Lys Leu Arg Val Gly Asp Arg Ile Val Thr Ile Cys Gly Thr
695 700 705
Ser Thr Glu Gly Met Thr His Thr Gln Ala Val Asn Leu Leu Lys
710 715 720
Asn Ala Ser Gly Ser Ile Glu Met Gln Val Val Ala Gly Gly Asp
725 730 735
Val Ser Val Val Thr Gly His Gln Gln Glu Pro Ala Ser Ser Ser
740 745 750
Leu Ser Phe Thr Gly Leu Thr Ser Ser Ser Ile Phe Gln Asp Asp
755 760 765
Leu Gly Pro Pro Gln Cys Lys Ser Ile Thr Leu Glu Arg Gly Pro
770 775 780
Asp Gly Leu Gly Phe Ser Ile Val Gly Gly Tyr Gly Ser Pro His
785 790 795
Gly Asp Leu Pro Ile Tyr Val Lys Thr Val Phe Ala Lys Gly Ala
800 805 810
Ala Ser Glu Asp Gly Arg Leu Lys Arg Gly Asp Gln Ile Ile Ala
815 820 825
Val Asn Gly Gln Ser Leu Glu Gly Val Thr His Glu Glu Ala Val
830 835 840

Ala Ile Leu Lys Arg Thr Lys Gly Thr Val Thr Leu Met Val Leu
845 850 855
Ser

<210> 13
<211> 361
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3800639CD1

<400> 13
Met Glu Thr Gly Ala Ala Glu Leu Tyr Asp Gln Ala Leu Leu Gly
1 5 10 15
Ile Leu Gln His Val Gly Asn Val Gln Asp Phe Leu Arg Val Leu
20 25 30
Phe Gly Phe Leu Tyr Arg Lys Thr Asp Phe Tyr Arg Leu Leu Arg
35 40 45
His Pro Ser Asp Arg Met Gly Phe Pro Pro Gly Ala Ala Gln Ala
50 55 60
Leu Val Leu Gln Val Phe Lys Thr Phe Asp His Met Ala Arg Gln
65 70 75
Asp Asp Glu Lys Arg Arg Gln Glu Leu Glu Glu Lys Ile Arg Arg
80 85 90
Lys Glu Glu Glu Glu Ala Lys Thr Val Ser Ala Ala Ala Ala Glu
95 100 105
Lys Glu Pro Val Pro Val Pro Val Gln Glu Ile Glu Ile Asp Ser
110 115 120
Thr Thr Glu Leu Asp Gly His Gln Glu Val Glu Lys Val Gln Pro
125 130 135
Pro Gly Pro Val Lys Glu Met Ala His Gly Ser Gln Glu Ala Glu
140 145 150
Ala Pro Gly Ala Val Ala Gly Ala Ala Glu Val Pro Arg Glu Pro
155 160 165
Pro Ile Leu Pro Arg Ile Gln Glu Gln Phe Gln Lys Asn Pro Asp
170 175 180
Ser Tyr Asn Gly Ala Val Arg Glu Asn Tyr Thr Trp Ser Gln Asp
185 190 195
Tyr Thr Asp Leu Glu Val Arg Val Pro Val Pro Lys His Val Val
200 205 210
Lys Gly Lys Gln Val Ser Val Ala Leu Ser Ser Ser Ser Ile Arg
215 220 225
Val Ala Met Leu Glu Glu Asn Gly Glu Arg Val Leu Met Glu Gly
230 235 240
Lys Leu Thr His Lys Ile Asn Thr Glu Ser Ser Leu Trp Ser Leu
245 250 255
Glu Pro Gly Lys Cys Val Leu Val Asn Leu Ser Lys Val Gly Glu
260 265 270
Tyr Trp Trp Asn Ala Ile Leu Glu Gly Glu Glu Pro Ile Asp Ile
275 280 285
Asp Lys Ile Asn Lys Glu Arg Ser Met Ala Thr Val Asp Glu Glu
290 295 300

```
Glu Gln Ala Val Leu Asp Arg Leu Thr Phe Asp Tyr His Gln Lys
                305                 310                 315
Leu Gln Gly Lys Pro Gln Ser His Glu Leu Lys Val His Glu Met
                320                 325                 330
Leu Lys Lys Gly Trp Asp Ala Glu Gly Ser Pro Phe Arg Gly Gln
                335                 340                 345
Arg Phe Asp Pro Ala Met Phe Asn Ile Ser Pro Gly Ala Val Gln
                350                 355                 360
Phe


<210> 14
<211> 632
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 533825CD1

<400> 14
Met Lys Ala Leu Leu Leu Leu Val Leu Pro Trp Leu Ser Pro Ala
  1               5                  10                  15
Asn Tyr Ile Asp Asn Val Gly Asn Leu His Phe Leu Tyr Ser Glu
                 20                  25                  30
Leu Cys Lys Gly Ala Ser His Tyr Gly Leu Thr Lys Asp Arg Lys
                 35                  40                  45
Arg Arg Ser Gln Asp Gly Cys Pro Asp Gly Cys Ala Ser Leu Thr
                 50                  55                  60
Ala Thr Ala Pro Ser Pro Glu Val Ser Ala Ala Ala Thr Ile Ser
                 65                  70                  75
Leu Met Thr Asp Glu Pro Gly Leu Asp Asn Pro Ala Tyr Val Ser
                 80                  85                  90
Ser Ala Glu Asp Gly Gln Pro Ala Ile Ser Pro Val Asp Ser Gly
                 95                 100                 105
Arg Ser Asn Arg Thr Arg Ala Arg Pro Phe Glu Arg Ser Thr Ile
                110                 115                 120
Arg Ser Arg Ser Phe Lys Lys Ile Asn Arg Ala Leu Ser Val Leu
                125                 130                 135
Arg Arg Thr Lys Ser Gly Ser Ala Val Ala Asn His Ala Asp Gln
                140                 145                 150
Gly Arg Glu Asn Ser Glu Asn Ile Thr Ala Pro Glu Val Phe Pro
                155                 160                 165
Arg Leu Tyr His Leu Ile Pro Asp Gly Glu Ile Thr Ser Ile Lys
                170                 175                 180
Ile Asn Arg Val Asp Pro Ser Glu Ser Leu Ser Ile Arg Leu Val
                185                 190                 195
Gly Gly Ser Glu Thr Pro Leu Val His Ile Ile Ile Gln His Ile
                200                 205                 210
Tyr Arg Asp Gly Val Ile Ala Arg Asp Gly Arg Leu Leu Pro Gly
                215                 220                 225
Asp Ile Ile Leu Lys Val Asn Gly Met Asp Ile Ser Asn Val Pro
                230                 235                 240
His Asn Tyr Ala Val Arg Leu Leu Arg Gln Pro Cys Gln Val Leu
                245                 250                 255
```

Trp Leu Thr Val Met Arg Glu Gln Lys Phe Arg Ser Arg Asn Asn
260 265 270
Gly Gln Ala Pro Asp Ala Tyr Arg Pro Arg Asp Asp Ser Phe His
275 280 285
Val Ile Leu Asn Lys Ser Ser Pro Glu Glu Gln Leu Gly Ile Lys
290 295 300
Leu Val Arg Lys Val Asp Glu Pro Gly Val Phe Ile Phe Asn Val
305 310 315
Leu Asp Gly Gly Val Ala Tyr Arg His Gly Gln Leu Glu Glu Asn
320 325 330
Asp Arg Val Leu Ala Ile Asn Gly His Asp Leu Arg Tyr Gly Ser
335 340 345
Pro Glu Ser Ala Ala His Leu Ile Gln Ala Ser Glu Arg Arg Val
350 355 360
His Leu Val Val Ser Arg Gln Val Arg Gln Arg Ser Pro Asp Ile
365 370 375
Phe Gln Glu Ala Gly Trp Asn Ser Asn Gly Ser Trp Ser Pro Gly
380 385 390
Pro Gly Glu Arg Ser Asn Thr Pro Lys Pro Leu His Pro Thr Ile
395 400 405
Thr Cys His Glu Lys Val Val Asn Ile Gln Lys Asp Pro Gly Glu
410 415 420
Ser Leu Gly Met Ala Val Ala Gly Gly Ala Ser His Arg Glu Trp
425 430 435
Asp Leu Pro Ile Tyr Val Ile Ser Val Glu Pro Gly Gly Val Ile
440 445 450
Ser Arg Asp Gly Arg Ile Lys Thr Gly Asp Ile Leu Leu Asn Val
455 460 465
Asp Gly Val Glu Leu Thr Glu Val Ser Arg Ser Glu Ala Val Ala
470 475 480
Leu Leu Lys Arg Thr Ser Ser Ser Ile Val Leu Lys Ala Leu Glu
485 490 495
Val Lys Glu Tyr Glu Pro Gln Glu Asp Cys Ser Ser Pro Ala Ala
500 505 510
Leu Asp Ser Asn His Asn Met Ala Pro Pro Ser Asp Trp Ser Pro
515 520 525
Ser Trp Val Met Trp Leu Glu Leu Pro Arg Cys Leu Tyr Asn Cys
530 535 540
Lys Asp Ile Val Leu Arg Arg Asn Thr Ala Gly Ser Leu Gly Phe
545 550 555
Cys Ile Val Gly Gly Tyr Glu Glu Tyr Asn Gly Asn Lys Pro Phe
560 565 570
Phe Ile Lys Ser Ile Val Glu Gly Thr Pro Ala Tyr Asn Asp Gly
575 580 585
Arg Ile Arg Cys Gly Asp Ile Leu Leu Ala Val Asn Gly Arg Ser
590 595 600
Thr Ser Gly Met Ile His Ala Cys Leu Ala Arg Leu Leu Lys Glu
605 610 615
Leu Lys Gly Arg Ile Thr Leu Thr Ile Val Ser Trp Pro Gly Thr
620 625 630
Phe Leu

<210> 15
<211> 391
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1311833CD1

<400> 15
Met Lys Met Lys Ile Gln Lys Lys Glu Lys Gln Leu Ser Asn Leu
1 5 10 15
Lys Val Leu Asn His Ser Pro Met Ser Asp Ala Ser Val Asn Phe
20 25 30
Asp Tyr Lys Ser Pro Ser Pro Phe Asp Cys Ser Thr Asp Gln Glu
35 40 45
Glu Lys Ile Glu Asp Val Ala Ser His Cys Leu Pro Gln Lys Asp
50 55 60
Leu Tyr Thr Ala Glu Glu Glu Ala Ala Thr Leu Phe Pro Arg Lys
65 70 75
Met Thr Ser His Asn Gly Met Glu Asp Ser Gly Gly Gly Gly Thr
80 85 90
Gly Val Lys Lys Lys Arg Lys Lys Lys Glu Pro Gly Asp Gln Glu
95 100 105
Gly Ala Ala Lys Gly Ser Lys Asp Arg Glu Pro Lys Pro Lys Arg
110 115 120
Lys Arg Glu Pro Lys Glu Pro Lys Glu Pro Arg Lys Ala Lys Glu
125 130 135
Pro Lys Lys Ala Lys Glu His Lys Glu Pro Lys Gln Lys Asp Gly
140 145 150
Ala Lys Lys Ala Arg Lys Pro Arg Glu Ala Ser Gly Thr Lys Glu
155 160 165
Ala Lys Glu Lys Arg Ser Cys Thr Asp Ser Ala Ala Arg Thr Lys
170 175 180
Ser Arg Lys Ala Ser Lys Glu Gln Gly Pro Thr Pro Val Glu Lys
185 190 195
Lys Lys Lys Gly Lys Arg Lys Ser Glu Thr Thr Val Glu Ser Leu
200 205 210
Glu Leu Asp Gln Gly Leu Thr Asn Pro Ser Leu Arg Ser Pro Glu
215 220 225
Glu Ser Thr Glu Ser Thr Asp Ser Gln Lys Arg Arg Ser Gly Arg
230 235 240
Gln Val Lys Arg Arg Lys Tyr Asn Glu Asp Leu Asp Phe Lys Val
245 250 255
Val Asp Asp Asp Gly Glu Thr Ile Ala Val Leu Gly Ala Gly Arg
260 265 270
Thr Ser Ala Leu Ser Ala Ser Thr Leu Ala Trp Gln Ala Glu Glu
275 280 285
Pro Pro Glu Asp Asp Ala Asn Ile Ile Glu Lys Ile Leu Ala Ser
290 295 300
Lys Thr Val Gln Glu Val His Pro Gly Glu Pro Pro Phe Asp Leu
305 310 315
Glu Leu Phe Tyr Val Lys Tyr Arg Asn Phe Ser Tyr Leu His Cys
320 325 330
Lys Trp Ala Thr Met Glu Glu Leu Glu Lys Asp Pro Arg Ile Ala

```
                    335                     340                     345
Gln Lys Ile Lys Arg Phe Arg Asn Lys Gln Ala Gln Met Lys His
                    350                     355                     360
Ile Phe Thr Glu Val Lys Gln Tyr Leu Leu Thr His Leu Thr Ala
                    365                     370                     375
Ala Phe Leu Ala Ala Val Asn Thr Val Phe Thr Phe Leu Ser Pro
                    380                     385                     390
Ser


<210> 16
<211> 490
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1342819CD1

<400> 16
Met Glu Asp Ser Ala Ser Ala Ser Leu Ser Ser Ala Ala Ala Thr
  1               5                  10                  15
Gly Thr Ser Thr Ser Thr Pro Ala Ala Pro Thr Ala Arg Lys Gln
                 20                  25                  30
Leu Asp Lys Glu Gln Val Arg Lys Ala Val Asp Ala Leu Leu Thr
                 35                  40                  45
His Cys Lys Ser Arg Lys Asn Asn Tyr Gly Leu Leu Leu Asn Glu
                 50                  55                  60
Asn Glu Ser Leu Phe Leu Met Val Val Leu Trp Lys Ile Pro Ser
                 65                  70                  75
Lys Glu Leu Arg Val Arg Leu Thr Leu Pro His Ser Ile Arg Ser
                 80                  85                  90
Asp Ser Glu Asp Ile Cys Leu Phe Thr Lys Asp Glu Pro Asn Ser
                 95                 100                 105
Thr Pro Glu Lys Thr Glu Gln Phe Tyr Arg Lys Leu Leu Asn Lys
                110                 115                 120
His Gly Ile Lys Thr Val Ser Gln Ile Ile Ser Leu Gln Thr Leu
                125                 130                 135
Lys Lys Glu Tyr Lys Ser Tyr Glu Ala Lys Leu Arg Leu Leu Ser
                140                 145                 150
Ser Phe Asp Phe Phe Leu Thr Asp Ala Arg Ile Arg Arg Leu Leu
                155                 160                 165
Pro Ser Leu Ile Gly Arg His Phe Tyr Gln Arg Lys Lys Val Pro
                170                 175                 180
Val Ser Val Asn Leu Leu Ser Lys Asn Leu Ser Arg Glu Ile Asn
                185                 190                 195
Asp Cys Ile Gly Gly Thr Val Leu Asn Ile Ser Lys Ser Gly Ser
                200                 205                 210
Cys Ser Ala Ile Arg Ile Gly His Val Gly Met Gln Ile Glu His
                215                 220                 225
Ile Ile Glu Asn Ile Val Ala Val Thr Lys Gly Leu Ser Glu Lys
                230                 235                 240
Leu Pro Glu Lys Trp Glu Ser Val Lys Leu Leu Phe Val Lys Thr
                245                 250                 255
Glu Lys Ser Ala Ala Leu Pro Ile Phe Ser Ser Phe Val Ser Asn
```

```
                    260                      265                      270
Trp Asp Glu Ala Thr Lys Arg Ser Leu Leu Asn Lys Lys Lys Lys
                    275                      280                      285
Glu Ala Arg Arg Lys Arg Arg Glu Arg Asn Phe Glu Lys Gln Lys
                    290                      295                      300
Glu Arg Lys Lys Lys Arg Gln Gln Ala Arg Lys Thr Ala Ser Val
                    305                      310                      315
Leu Ser Lys Asp Asp Val Ala Pro Glu Ser Gly Asp Thr Thr Val
                    320                      325                      330
Lys Lys Pro Glu Ser Lys Lys Glu Gln Thr Pro Glu His Gly Lys
                    335                      340                      345
Lys Lys Arg Gly Arg Gly Lys Ala Gln Val Lys Ala Thr Asn Glu
                    350                      355                      360
Ser Glu Asp Glu Ile Pro Gln Leu Val Pro Ile Gly Lys Lys Thr
                    365                      370                      375
Pro Ala Asn Glu Lys Val Glu Ile Gln Lys His Ala Thr Gly Lys
                    380                      385                      390
Lys Ser Pro Ala Lys Ser Pro Asn Pro Ser Thr Pro Arg Gly Lys
                    395                      400                      405
Lys Arg Lys Ala Leu Pro Ala Ser Glu Thr Pro Lys Ala Ala Glu
                    410                      415                      420
Ser Glu Thr Pro Gly Lys Ser Pro Glu Lys Lys Pro Lys Ile Lys
                    425                      430                      435
Glu Glu Ala Val Lys Glu Lys Ser Pro Ser Leu Gly Lys Lys Asp
                    440                      445                      450
Ala Arg Gln Thr Pro Lys Lys Pro Glu Ala Lys Phe Phe Thr Thr
                    455                      460                      465
Pro Ser Lys Ser Val Arg Lys Ala Ser His Thr Pro Lys Lys Trp
                    470                      475                      480
Pro Lys Lys Pro Lys Val Pro Gln Ser Thr
                    485                      490
```

<210> 17
<211> 252
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1871288CD1

<400> 17

```
Met Ala Glu Leu Glu Phe Val Gln Ile Ile Ile Ile Val Val Val
  1                 5                       10                       15
Met Met Val Met Val Val Val Ile Thr Cys Leu Leu Ser His Tyr
                   20                       25                       30
Lys Leu Ser Ala Arg Ser Phe Ile Ser Arg His Ser Gln Gly Arg
                   35                       40                       45
Arg Arg Glu Asp Ala Leu Ser Ser Glu Gly Cys Leu Trp Pro Ser
                   50                       55                       60
Glu Ser Thr Val Ser Gly Asn Gly Ile Pro Glu Pro Gln Val Tyr
                   65                       70                       75
Ala Pro Pro Arg Pro Thr Asp Arg Leu Ala Val Pro Pro Phe Ala
                   80                       85                       90
```

```
Gln Arg Glu Arg Phe His Arg Phe Gln Pro Thr Tyr Pro Tyr Leu
                95                  100                 105
Gln His Glu Ile Asp Leu Pro Pro Thr Ile Ser Leu Ser Asp Gly
                110                 115                 120
Glu Glu Pro Pro Pro Tyr Gln Gly Pro Cys Thr Leu Gln Leu Arg
                125                 130                 135
Asp Pro Glu Gln Gln Leu Glu Leu Asn Arg Glu Ser Val Arg Ala
                140                 145                 150
Pro Pro Asn Arg Thr Ile Phe Asp Ser Asp Leu Met Asp Ser Ala
                155                 160                 165
Arg Leu Gly Gly Pro Cys Pro Pro Ser Ser Asn Ser Gly Ile Ser
                170                 175                 180
Ala Thr Cys Tyr Gly Ser Gly Gly Arg Met Glu Gly Pro Pro Pro
                185                 190                 195
Thr Tyr Ser Glu Val Ile Gly His Tyr Pro Gly Ser Ser Phe Gln
                200                 205                 210
His Gln Gln Ser Ser Gly Pro Pro Ser Leu Leu Glu Gly Thr Arg
                215                 220                 225
Leu His His Thr His Ile Ala Pro Leu Glu Ser Ala Ala Ile Trp
                230                 235                 240
Ser Lys Glu Lys Asp Lys Gln Lys Gly His Pro Leu
                245                 250


<210> 18
<211> 142
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2587338CD1

<400> 18
Met Glu Ser Ala Arg Glu Asn Ile Asp Leu Gln Pro Gly Ser Ser
 1              5                   10                  15
Asp Pro Arg Ser Gln Pro Ile Asn Leu Asn His Tyr Ala Thr Lys
                20                  25                  30
Lys Ser Val Ala Glu Ser Met Leu Asp Val Ala Leu Phe Met Ser
                35                  40                  45
Asn Ala Met Arg Leu Lys Ala Val Leu Glu Gln Gly Pro Ser Ser
                50                  55                  60
His Tyr Tyr Thr Thr Leu Val Thr Leu Ile Ser Leu Ser Leu Leu
                65                  70                  75
Leu Gln Val Val Ile Gly Val Leu Leu Val Val Ile Ala Arg Leu
                80                  85                  90
Asn Leu Asn Glu Val Glu Lys Gln Trp Arg Leu Asn Gln Leu Asn
                95                  100                 105
Asn Gly Ser His Ile Leu Val Phe Phe Thr Val Val Ile Asn Gly
                110                 115                 120
Phe Ile Thr Gly Phe Gly Ala His Lys Thr Arg Val Leu Ala Cys
                125                 130                 135
Gln Asp Ser Arg Asn Pro Leu
                140
```

<210> 19
<211> 67
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2821211CD1

<400> 19
Met Glu Ile Ile Glu Asn Ser Phe His Ile Asn Gly Leu Lys Ile
1 5 10 15
Asn Gln Arg Thr Leu Cys Val His Val Cys Ile Ser Ala His Arg
20 25 30
Asn Ile Tyr Thr Tyr Val Asp Tyr Ile His Val Cys Ile Tyr Val
35 40 45
Tyr Ile Tyr Ile His Leu Tyr Lys Cys Ile Tyr Thr Tyr Thr Tyr
50 55 60
Asn Val Cys Met Cys Ile Tyr
65

<210> 20
<211> 455
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2824832CD1

<400> 20
Met Phe Gln Phe His Ala Gly Ser Trp Glu Ser Trp Cys Cys Cys
1 5 10 15
Cys Leu Ile Pro Ala Asp Arg Pro Trp Asp Arg Gly Gln His Trp
20 25 30
Gln Leu Glu Met Ala Asp Thr Arg Ser Val His Glu Thr Arg Phe
35 40 45
Glu Ala Ala Val Lys Val Ile Gln Ser Leu Pro Lys Asn Gly Ser
50 55 60
Phe Gln Pro Thr Asn Glu Met Met Leu Lys Phe Tyr Ser Phe Tyr
65 70 75
Lys Gln Ala Thr Glu Gly Pro Cys Lys Leu Ser Arg Pro Gly Phe
80 85 90
Trp Asp Pro Ile Gly Arg Tyr Lys Trp Asp Ala Trp Ser Ser Leu
95 100 105
Gly Asp Met Thr Lys Glu Glu Ala Met Ile Ala Tyr Val Glu Glu
110 115 120
Met Lys Lys Ile Ile Glu Thr Met Pro Met Thr Glu Lys Val Glu
125 130 135
Glu Leu Leu Arg Val Ile Gly Pro Phe Tyr Glu Ile Val Glu Asp
140 145 150
Lys Lys Ser Gly Arg Ser Ser Asp Ile Thr Ser Asp Leu Gly Asn

```
                    155                     160                     165
Val Leu Thr Ser Thr Pro Asn Ala Lys Thr Val Asn Gly Lys Ala
                    170                     175                     180
Glu Ser Ser Asp Ser Gly Ala Glu Ser Glu Glu Glu Glu Ala Gln
                    185                     190                     195
Glu Glu Val Lys Gly Ala Glu Gln Ser Asp Asn Asp Ile Asn Asp
                    200                     205                     210
Asp His Val Glu Asp Val Thr Gly Ile Gln His Leu Thr Ser Asp
                    215                     220                     225
Ser Asp Ser Glu Val Tyr Cys Asp Ser Met Glu Gln Phe Gly Gln
                    230                     235                     240
Glu Glu Ser Leu Asp Ser Phe Thr Ser Asn Asn Gly Pro Phe Gln
                    245                     250                     255
Tyr Tyr Leu Gly Gly His Ser Ser Gln Pro Met Glu Asn Ser Gly
                    260                     265                     270
Phe Arg Glu Asp Ile Gln Val Pro Pro Gly Asn Gly Asn Ile Gly
                    275                     280                     285
Asn Met Gln Val Val Ala Val Glu Gly Lys Gly Glu Val Lys His
                    290                     295                     300
Gly Gly Glu Asp Gly Arg Asn Asn Ser Gly Ala Pro His Arg Glu
                    305                     310                     315
Lys Arg Gly Gly Glu Thr Asp Glu Phe Ser Asn Val Arg Arg Gly
                    320                     325                     330
Arg Gly His Arg Met Gln His Leu Ser Glu Gly Thr Lys Gly Arg
                    335                     340                     345
Gln Val Gly Ser Gly Gly Asp Gly Glu Arg Trp Gly Ser Asp Arg
                    350                     355                     360
Gly Ser Arg Gly Ser Leu Asn Glu Gln Ile Ala Leu Val Leu Met
                    365                     370                     375
Arg Leu Gln Glu Asp Met Gln Asn Val Leu Gln Arg Leu Gln Lys
                    380                     385                     390
Leu Glu Thr Leu Thr Ala Leu Gln Ala Lys Ser Ser Thr Ser Thr
                    395                     400                     405
Leu Gln Thr Ala Pro Gln Pro Thr Ser Gln Arg Pro Ser Trp Trp
                    410                     415                     420
Pro Phe Glu Met Ser Pro Gly Val Leu Thr Phe Ala Ile Ile Trp
                    425                     430                     435
Pro Phe Ile Ala Gln Trp Leu Val Tyr Leu Tyr Tyr Gln Arg Arg
                    440                     445                     450
Arg Arg Lys Leu Asn
                    455


<210> 21
<211> 252
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3070147CD1

<400> 21
Met Gln Leu Thr Arg Cys Cys Phe Val Phe Leu Val Gln Gly Ser
  1               5                  10                  15
```

Leu Tyr Leu Val Ile Cys Gly Gln Asp Asp Gly Pro Pro Gly Ser
20 25 30
Glu Asp Pro Glu Arg Asp Asp His Glu Gly Gln Pro Arg Pro Arg
35 40 45
Val Pro Arg Lys Arg Gly His Ile Ser Pro Lys Ser Arg Pro Met
50 55 60
Ala Asn Ser Thr Leu Leu Gly Leu Leu Ala Pro Thr Gly Glu Ala
65 70 75
Trp Gly Ile Leu Gly Gln Pro Pro Asn Arg Pro Asn His Ser Pro
80 85 90
Pro Pro Ser Ala Lys Val Lys Lys Ile Phe Gly Trp Gly Asp Phe
95 100 105
Tyr Ser Asn Ile Lys Thr Val Ala Leu Asn Leu Leu Val Thr Gly
110 115 120
Lys Ile Val Asp His Gly Asn Gly Thr Phe Ser Val His Phe Gln
125 130 135
His Asn Ala Thr Gly Gln Gly Asn Ile Ser Ile Ser Leu Val Pro
140 145 150
Pro Ser Lys Ala Val Glu Phe His Gln Glu Gln Gln Ile Phe Ile
155 160 165
Glu Ala Lys Ala Ser Lys Ile Phe Asn Cys Arg Met Glu Trp Glu
170 175 180
Lys Val Glu Arg Gly Arg Arg Thr Ser Leu Cys Thr His Asp Pro
185 190 195
Ala Lys Ile Cys Ser Arg Asp His Ala Gln Ser Ser Ala Thr Trp
200 205 210
Ser Cys Ser Gln Pro Phe Lys Val Val Cys Val Tyr Ile Ala Phe
215 220 225
Tyr Ser Thr Asp Tyr Arg Leu Val Gln Lys Val Cys Pro Asp Tyr
230 235 240
Asn Tyr His Ser Asp Thr Pro Tyr Tyr Pro Ser Gly
245 250

<210> 22
<211> 149
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3271841CD1

<400> 22
Met Glu Ser Arg Gly Lys Ser Ala Ser Ser Pro Lys Pro Asp Thr
1 5 10 15
Lys Val Pro Gln Val Thr Thr Glu Ala Lys Val Pro Pro Ala Ala
20 25 30
Asp Gly Lys Ala Pro Leu Thr Lys Pro Ser Lys Lys Glu Ala Pro
35 40 45
Ala Glu Lys Gln Gln Pro Pro Ala Ala Pro Thr Thr Ala Pro Ala
50 55 60
Lys Lys Thr Ser Ala Lys Ala Asp Pro Ala Leu Leu Asn Asn His
65 70 75
Ser Asn Leu Lys Pro Ala Pro Thr Val Pro Ser Ser Pro Asp Ala

```
                    80                      85                      90
Thr Pro Glu Pro Lys Gly Pro Gly Asp Gly Ala Glu Glu Asp Glu
                    95                      100                     105
Ala Ala Ser Gly Gly Pro Gly Gly Arg Gly Pro Trp Ser Cys Glu
                    110                     115                     120
Asn Phe Asn Pro Leu Leu Val Ala Gly Gly Val Ala Val Ala Ala
                    125                     130                     135
Ile Ala Leu Ile Leu Gly Val Ala Phe Leu Val Arg Lys Lys
                    140                     145
```

<210> 23
<211> 204
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3537827CD1

<400> 23

```
Met Met Pro Ser Cys Asn Arg Ser Cys Ser Cys Ser Arg Gly Pro
  1                 5                       10                      15
Ser Val Glu Asp Gly Lys Trp Tyr Gly Val Arg Ser Tyr Leu His
                    20                      25                      30
Leu Phe Tyr Glu Asp Cys Ala Gly Thr Ala Leu Ser Asp Asp Pro
                    35                      40                      45
Glu Gly Pro Pro Val Leu Cys Pro Arg Arg Pro Trp Pro Ser Leu
                    50                      55                      60
Cys Trp Lys Ile Ser Leu Ser Ser Gly Thr Leu Leu Leu Leu Leu
                    65                      70                      75
Gly Val Ala Ala Leu Thr Thr Gly Tyr Ala Val Pro Pro Lys Leu
                    80                      85                      90
Glu Gly Ile Gly Glu Gly Glu Phe Leu Val Leu Asp Gln Arg Ala
                    95                      100                     105
Ala Asp Tyr Asn Gln Ala Leu Gly Thr Cys Arg Leu Ala Gly Thr
                    110                     115                     120
Ala Leu Cys Val Ala Ala Gly Val Leu Leu Ala Ile Cys Leu Phe
                    125                     130                     135
Trp Ala Met Ile Gly Trp Leu Ser Gln Asp Thr Lys Ala Glu Pro
                    140                     145                     150
Leu Asp Pro Glu Ala Asp Ser His Val Glu Val Phe Gly Asp Glu
                    155                     160                     165
Pro Glu Gln Gln Leu Ser Pro Ile Phe Arg Asn Ala Ser Gly Gln
                    170                     175                     180
Ser Trp Phe Ser Pro Pro Ala Ser Pro Phe Gly Gln Ser Ser Val
                    185                     190                     195
Gln Thr Ile Gln Pro Lys Arg Asp Ser
                    200
```

<210> 24
<211> 367
<212> PRT
<213> Homo sapiens

```
<220>
<221> misc_feature
<223> Incyte ID No.: 3729267CD1

<400> 24
Met Ala Ser Glu Leu Cys Lys Thr Ile Ser Val Ala Arg Leu Glu
 1               5                  10                  15
Lys His Lys Asn Leu Phe Leu Asn Tyr Arg Asn Leu His His Phe
                20                  25                  30
Pro Leu Glu Leu Leu Lys Asp Glu Gly Leu Gln Tyr Leu Glu Arg
                35                  40                  45
Leu Tyr Met Lys Arg Asn Ser Leu Thr Ser Leu Pro Glu Asn Leu
                50                  55                  60
Ala Gln Lys Leu Pro Asn Leu Val Glu Leu Tyr Leu His Ser Asn
                65                  70                  75
Asn Ile Val Val Val Pro Glu Ala Ile Gly Ser Leu Val Lys Leu
                80                  85                  90
Gln Cys Leu Asp Leu Ser Asp Asn Ala Leu Glu Ile Val Cys Pro
                95                 100                 105
Glu Ile Gly Arg Leu Arg Ala Leu Arg His Leu Arg Leu Ala Asn
               110                 115                 120
Asn Gln Leu Gln Phe Leu Pro Pro Glu Val Gly Asp Leu Lys Glu
               125                 130                 135
Leu Gln Thr Leu Asp Ile Ser Thr Asn Arg Leu Leu Thr Leu Pro
               140                 145                 150
Glu Arg Leu His Met Cys Leu Ser Leu Gln Tyr Leu Thr Val Asp
               155                 160                 165
Arg Asn Arg Leu Trp Tyr Val Pro Arg His Leu Cys Gln Leu Pro
               170                 175                 180
Ser Leu Asn Glu Leu Ser Met Ala Gly Asn Arg Leu Ala Phe Leu
               185                 190                 195
Pro Leu Asp Leu Gly Arg Ser Arg Glu Leu Gln Tyr Val Tyr Val
               200                 205                 210
Asp Asn Asn Ile His Leu Lys Gly Leu Pro Ser Tyr Leu Tyr Asn
               215                 220                 225
Lys Val Ile Gly Cys Ser Gly Cys Gly Ala Pro Ile Gln Val Ser
               230                 235                 240
Glu Val Lys Leu Leu Ser Phe Ser Ser Gly Gln Arg Thr Val Phe
               245                 250                 255
Leu Pro Ala Glu Val Lys Ala Ile Gly Thr Glu His Asp His Val
               260                 265                 270
Leu Pro Leu Gln Glu Leu Ala Met Arg Gly Leu Tyr His Thr Tyr
               275                 280                 285
His Ser Leu Leu Lys Asp Leu Asn Phe Leu Ser Pro Ile Ser Leu
               290                 295                 300
Pro Arg Ser Leu Leu Glu Leu Leu His Cys Pro Leu Gly His Cys
               305                 310                 315
His Arg Cys Ser Glu Pro Met Phe Thr Ile Val Tyr Pro Lys Leu
               320                 325                 330
Phe Pro Leu Arg Glu Thr Pro Met Ala Gly Leu His Gln Trp Lys
               335                 340                 345
Thr Thr Val Ser Phe Val Ala Tyr Cys Cys Ser Thr Gln Cys Leu
               350                 355                 360
Gln Thr Phe Asp Leu Leu Ser
               365
```

<210> 25
<211> 681
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3768771CD1

<400> 25
Met Cys Thr Tyr Ile Asn Met Glu Asn Phe Thr Leu Ala Arg Asp
1 5 10 15
Glu Lys Gly Asn Val Leu Leu Glu Asp Gly Lys Gly Arg Cys Pro
20 25 30
Phe Asp Pro Asn Phe Lys Ser Thr Ala Leu Val Val Asp Gly Glu
35 40 45
Leu Tyr Thr Gly Thr Val Ser Ser Phe Gln Gly Asn Asp Pro Ala
50 55 60
Ile Ser Arg Ser Gln Ser Leu Arg Pro Thr Lys Thr Glu Ser Ser
65 70 75
Leu Asn Trp Leu Gln Asp Pro Ala Phe Val Ala Ser Ala Tyr Ile
80 85 90
Pro Glu Ser Leu Gly Ser Leu Gln Gly Asp Asp Asp Lys Ile Tyr
95 100 105
Phe Phe Phe Ser Glu Thr Gly Gln Glu Phe Glu Phe Phe Glu Asn
110 115 120
Thr Ile Val Ser Arg Ile Ala Arg Ile Cys Lys Gly Asp Glu Gly
125 130 135
Gly Glu Arg Val Leu Gln Gln Arg Trp Thr Ser Phe Leu Lys Ala
140 145 150
Gln Leu Leu Cys Ser Arg Pro Asp Asp Gly Phe Pro Phe Asn Val
155 160 165
Leu Gln Asp Val Phe Thr Leu Ser Pro Ser Pro Gln Asp Trp Arg
170 175 180
Asp Thr Leu Phe Tyr Gly Val Phe Thr Ser Gln Trp His Arg Gly
185 190 195
Thr Thr Glu Gly Ser Ala Val Cys Val Phe Thr Met Lys Asp Val
200 205 210
Gln Arg Val Phe Ser Gly Leu Tyr Lys Glu Val Asn Arg Glu Thr
215 220 225
Gln Gln Trp Tyr Thr Val Thr His Pro Val Pro Thr Pro Arg Pro
230 235 240
Gly Ala Cys Ile Thr Asn Ser Ala Arg Glu Arg Lys Ile Asn Ser
245 250 255
Ser Leu Gln Leu Pro Asp Arg Val Leu Asn Phe Leu Lys Asp His
260 265 270
Phe Leu Met Asp Gly Gln Val Arg Ser Arg Met Leu Leu Leu Gln
275 280 285
Pro Gln Ala Arg Tyr Gln Arg Val Ala Val His Arg Val Pro Gly
290 295 300
Leu His His Thr Tyr Asp Val Leu Phe Leu Gly Thr Gly Asp Gly
305 310 315

```
Arg Leu His Lys Ala Val Ser Val Gly Pro Arg Val His Ile Ile
                320                 325                 330
Glu Glu Leu Gln Ile Phe Ser Ser Gly Gln Pro Val Gln Asn Leu
                335                 340                 345
Leu Leu Asp Thr His Arg Gly Leu Leu Tyr Ala Ala Ser His Ser
                350                 355                 360
Gly Val Val Gln Val Pro Met Ala Asn Cys Ser Leu Tyr Arg Ser
                365                 370                 375
Cys Gly Asp Cys Leu Leu Ala Arg Asp Pro Tyr Cys Ala Trp Ser
                380                 385                 390
Gly Ser Ser Cys Lys His Val Ser Leu Tyr Gln Pro Gln Leu Ala
                395                 400                 405
Thr Arg Pro Trp Ile Gln Asp Ile Glu Gly Ala Ser Ala Lys Asp
                410                 415                 420
Leu Cys Ser Ala Ser Ser Val Val Ser Pro Ser Phe Val Pro Thr
                425                 430                 435
Gly Glu Lys Pro Cys Glu Gln Val Gln Phe Gln Pro Asn Thr Val
                440                 445                 450
Asn Thr Leu Ala Cys Pro Leu Leu Ser Asn Leu Ala Thr Arg Leu
                455                 460                 465
Trp Leu Arg Asn Gly Ala Pro Val Asn Ala Ser Ala Ser Cys His
                470                 475                 480
Val Leu Pro Thr Gly Asp Leu Leu Leu Val Gly Thr Gln Gln Leu
                485                 490                 495
Gly Glu Phe Gln Cys Trp Ser Leu Glu Glu Gly Phe Gln Gln Leu
                500                 505                 510
Val Ala Ser Tyr Cys Pro Glu Val Val Glu Asp Gly Val Ala Asp
                515                 520                 525
Gln Thr Asp Glu Gly Gly Ser Val Pro Val Ile Ile Ser Thr Ser
                530                 535                 540
Arg Val Ser Ala Pro Ala Gly Gly Lys Ala Ser Trp Gly Ala Asp
                545                 550                 555
Arg Ser Tyr Trp Lys Glu Phe Leu Val Met Cys Thr Leu Phe Val
                560                 565                 570
Leu Ala Val Leu Leu Pro Val Leu Phe Leu Leu Tyr Arg His Arg
                575                 580                 585
Asn Ser Met Lys Val Phe Leu Lys Gln Gly Glu Cys Ala Ser Val
                590                 595                 600
His Pro Lys Thr Cys Pro Val Val Leu Pro Pro Glu Thr Arg Pro
                605                 610                 615
Leu Asn Gly Leu Gly Pro Pro Ser Thr Pro Leu Asp His Arg Gly
                620                 625                 630
Tyr Gln Ser Leu Ser Asp Ser Pro Pro Gly Ser Arg Val Phe Thr
                635                 640                 645
Glu Ser Glu Lys Arg Pro Leu Ser Ile Gln Asp Ser Phe Val Glu
                650                 655                 660
Val Ser Pro Val Cys Pro Arg Pro Arg Val Arg Leu Gly Ser Glu
                665                 670                 675
Ile Arg Asp Ser Val Val
                680
```

<210> 26
<211> 137
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 4248993CD1

<400> 26
Met Gly Arg Lys Leu Asp Leu Ser Gly Leu Thr Asp Asp Glu Thr
1 5 10 15
Glu His Val Leu Gln Val Val Gln Arg Asp Phe Asn Leu Arg Lys
20 25 30
Lys Glu Glu Glu Arg Leu Ser Glu Leu Lys Gln Lys Leu Asp Glu
35 40 45
Glu Gly Ser Lys Cys Ser Ile Leu Ser Lys His Gln Gln Phe Val
50 55 60
Glu His Cys Cys Met Arg Cys Cys Ser Pro Phe Thr Phe Leu Val
65 70 75
Asn Thr Lys Arg Gln Cys Gly Asp Cys Lys Phe Asn Val Cys Lys
80 85 90
Ser Cys Cys Ser Tyr Gln Lys His Glu Lys Ala Trp Val Cys Cys
95 100 105
Val Cys Gln Gln Ala Arg Leu Leu Arg Ala Gln Ser Leu Glu Trp
110 115 120
Phe Tyr Asn Asn Val Lys Ser Arg Phe Lys Arg Phe Gly Ser Ala
125 130 135
Arg Phe

<210> 27
<211> 117
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 5402418CD1

<400> 27
Met Lys Phe Gln Tyr Lys Glu Asp His Pro Phe Glu Tyr Arg Lys
1 5 10 15
Lys Glu Gly Glu Lys Ile Arg Lys Lys Tyr Pro Asp Arg Val Pro
20 25 30
Val Ile Val Glu Lys Ala Pro Lys Ala Arg Val Pro Asp Leu Asp
35 40 45
Lys Arg Lys Tyr Leu Val Pro Ser Asp Leu Thr Val Gly Gln Phe
50 55 60
Tyr Phe Leu Ile Arg Lys Arg Ile His Leu Arg Pro Glu Asp Ala
65 70 75
Leu Phe Phe Phe Val Asn Asn Thr Ile Pro Pro Thr Ser Ala Thr
80 85 90

```
Met Gly Gln Leu Tyr Glu Asp Asn His Glu Glu Asp Tyr Phe Leu
                 95                  100                 105
Tyr Val Ala Tyr Ser Asp Glu Ser Val Tyr Gly Lys
                110                 115
```

<210> 28
<211> 1058
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2417014CB1

<400> 28

```
cgagatcgca gcccaaccca tggccgggtc tcctagccgc gccgcgggcc ggcgactgca   60
gcttcccctg ctgtgcctct tcctccaggg cgccactgcc gtcctctttg ctgtctttgt  120
ccgctacaac cacaaaaccg acgctgccct ctggcaccgg agcaaccaca gtaacgcgga  180
caatgaattt tactttcgct acccaaaaga gtctcactct gttgcccagg ctggagtgca  240
acgacgcaat ctcggctcac tgcaaccttc acctcccaga tggagtttcg ctcttgttgc  300
ccaggctgga gtgcaatggc acaatctcgg ctcaccacaa cctctgcctc ccgggttcaa  360
gcgattctcc tgcctcagtc tcctgagtag ctgggattac agcctggaga gtgtgtttcc  420
actcatagcc gagggccagc gcagtgccac gtcacaggcc atgcaccagc tcttcgggct  480
gtttgtcaca ctgatgtttg cctctgtggg cgggggcctt ggagggctcc tgctgaagct  540
accctttctg gactcccccc ccagactccc agcactacga ggaccaagtt cactggcagg  600
tgcctggcga gcatgaggat aaagcccaga gacctctgag ggtggaggag gcagacactc  660
aggcctaacc cactgccagc ccctgagagg acacgctcct tttcgaagat gctgactggc  720
tgctactagg aagttctttt tgagctccca ttcctccagc tgcaagaagg gagccatgag  780
ccagaaggag gccccttcc acaggcagcg tctccacagg gagaggggca acaggaggct  840
gggaaatggt ggggagtggg gccgtaactg ggtacaatag ggggaacctc accagatgcc  900
caacccgact gccctaccag cctgcacatg ggtagaagag gccaaattga ggcacccaag  960
tgatccactg gccccacgtc acacagttac agtgaagccc aagccaggcc tggttgaggg 1020
tgataaacgc cactgtgcgg caccgcaaaa aaaaaaaa                         1058
```

<210> 29
<211> 2235
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2634931CB1

<400> 29

```
cggccacccg tccgaccaca ccagggcaac tgtagtgcca gtgcctggtt ccacccgggg   60
ggcatctgag aactgtgtcc ttccattcct gagtccagca cttcccaggc caggaactca  120
cacagctttt ggcctgagcc cccgttacca agagaaagga ggtttttgcc aaggactcca  180
aggggagtgc acttgatgct ggtcgggacc caaagcaccc agccctccct gagacattgt  240
gtgagtcggg ctgggcctca aacacggccc ccactgcccc accccagcca gggtggtgct  300
tgtgtgggta ggactttaaa tccagctgcc agacccctgg acgggagaag gagagacggc  360
tggccaccat gcacggctcc tgcagtttcc tgatgcttct gctgccgcta ctgctactgc  420
tggtggccac cacaggcccc gttggagccc tcacagatga ggagaaacgt ttgatggtgg  480
agctgcacaa cctctaccgg gcccaggtat ccccgacggc ctcagacatg ctgcacatga  540
```

gatgggacga ggagctggcc gccttcgcca aggcctacgc acggcagtgc gtgtggggcc 600
acaacaagga gcgcgggcgc cgcggcgaga atctgttcgc catcacagac gagggcatgg 660
acgtgccgct ggccatggag gagtggcacc acgagcgtga gcactacaac ctcagcgccg 720
ccacctgcag cccaggccag atgtgcggcc actacacgca ggtggtatgg gccaagacag 780
agaggatcgg ctgtggttcc cacttctgtg agaagctcca gggtgttgag gagaccaaca 840
tcgaattact ggtgtgcaac tatgagcctc cggggaacgt gaaggggaaa cggccctacc 900
aggaggggac tccgtgctcc caatgtccct ctggctacca ctgcaagaac tccctctgtg 960
aacccatcgg aagcccggaa gatgctcagg atttgcctta cctggtaact gaggccccat 1020
ccttccgggc gactgaagca tcagactcta ggaaaatggg tactccttct tccctagcaa 1080
cggggattcc ggctttcttg gtaacagagg tctcaggctc cctggcaacc aaggctctgc 1140
ctgctgtgga aacccaggcc ccaacttcct tagcaacgaa agacccgccc tccatggcaa 1200
cagaggctcc accttgcgta acaactgagg tcccttccat tttggcagct cacagcctgc 1260
cctccttgga tgaggagcca gttaccttcc ccaaatcgac ccatgttcct atcccaaaat 1320
cagcagacaa agtgacagac aaaacaaaag tgccctctag gagcccagag aactctctgg 1380
accccaagat gtccctgaca ggggcaaggg agctcctacc ccatgcccag gaggaggctg 1440
aggctgaggc tgagttgcct ccttccagtg aggtcttggc ctcagttttt ccagcccagg 1500
acaagccagg tgagctgcag gccacactgg accacacggg gcacacctcc tccaagtccc 1560
tgcccaattt ccccaatacc tctgccaccg ctaatgccac gggtgggcgt gccctggctc 1620
tgcagtcgtc cttgccaggt gcagagggcc ctgacaagcc tagcgtcgtg tcagggctga 1680
actcgggccc tggtcatgtg tggggccctc tcctgggact actgctcctg cctcctctgg 1740
tgttggctgg aatcttctga aggggatacc actcaaaggg tgaagaggtc agctgtcctc 1800
ctgtcatctt ccccaccctg tccccagccc ctaaacaaga tacttcttgg ttaaggccct 1860
ccggaaggga aaggctacgg ggcatgtgcc tcatcacacc atccatcctg gaggcacaag 1920
gcctggctgg ctgcgagctc agggggccgc ctgaggactg cacaccgggc ccacacctct 1980
cctgcccctc cctcctgagt cctgggggtg ggaggatttg agggagctca ctgcctacct 2040
ggcctggggc tgtctgccca cacagcatgt gcgctctccc tgagtgcctg tgtagctggg 2100
gatggggatt cctaggggca gatgaaggac aagccccact ggagtggggt tctttgagtg 2160
ggggaggcag ggacgaggga aggaaagcaa ctcctgactc tccaataaaa acctgtccaa 2220
cctgtgaaaa aaaaa 2235

<210> 30
<211> 1559
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 110960CB1

<400> 30
cccaggcccg gccctctcct ccccccgcgc cgatggtacg cgccggctcg cctggcccgg 60
ctgcagtgga tgttgctgga acccacgcgg aggaaggaag agacgcaggc aggctgcggt 120
tacccaagcg gccacccggg cctcagggac cccttccccg agagacggca ccatgaccca 180
gggaaagctc tccgtggcta acaaggcccc tgggaccgag gggcagcagc aggtgcatgg 240
cgagaagaag gaggctccag cagtgccctc agccccaccc tcctatgagg aagccacctc 300
tggggagggg atgaaggcag gggccttccc cccagccccc acagcggtgc ctctccaccc 360
tagctgggcc tatgtggacc ccagcagcag ctccagctat gacaacggtt tccccaccgg 420
agaccatgag ctcttcacca ctttcagctg ggatgaccag aaagttcgtc gagtctttgt 480
cagaaaggtc tacaccatcc tgctgattca gctgctggtg accttggctg tcgtggctct 540
ctttactttc tgtgaccctg tcaaggacta tgtccaggcc aacccaggct ggtactgggc 600
atcctatgct gtgttctttg caacctacct gaccctggct tgctgttctg gacccaggag 660
gcatttcccc tggaacctga ttctcctgac cgtctttacc ctgtccatgg cctacctcac 720
tgggatgctg tccagctact acaacaccac ctccgtgctg ctgtgcctgg gcatcacggc 780
ccttgtctgc ctctcagtca ccgtcttcag cttccagacc aagttcgact tcacctcctg 840

ccagggcgtg ctcttcgtgc ttctcatgac tcttttcttc agcggactca tcctggccat 900
cctcctaccc ttccaatatg tgccctggct ccatgcagtt tatgcagcac tgggagcggg 960
tgtatttaca ttgttcctgg cacttgacac ccagttgctg atgggtaacc gacgccactc 1020
gctgagccct gaggagtata tttttggagc cctcaacatt tacctagaca tcatctatat 1080
cttcaccttc ttcctgcagc tttttggcac taaccgagaa tgaggagccc tccctgcccc 1140
accgtcctcc agagaatgcg cccctcctgg ttccctgtcc ctcccctgcg ctcctgcgag 1200
accagatata aaactagctg ccaacccagc ctgtggccag gtcactgtct accccagccc 1260
agcccagccc tctgccgctt gtacatacgc catggggacc ctgaggaact gaggccacgt 1320
caatccctgt gccgccccat tcgcccgtta catcttccaa actgggacgg tcaaggctga 1380
aggctcctct gggtttgagg gtccaaggga caaggaggag aagcctagca ggatttcaga 1440
tgcaggagag agacccaggg aagcccggca gagccctgag ccccactgca attctcctag 1500
ggctgcacat catgtggctt agggacactg tctgcatcca gtctgtgtct cctgtcttc 1559

<210> 31
<211> 876
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 380721CB1

<400> 31
cccacgcgtc cgcccacgcg tccgattttt catctttttt caggttgacg atgtgtcaca 60
ctgtgtaagg gaatcgcatg gagatgggca ttccgaactg ttaatgggga catgggactc 120
cagttgtctc tgatcacttg tgtggatttt cctggcgtag aacgacagaa gccgctagta 180
agtcgccaag acctacagca ggaattctgc accaaagggc ataaaatctt gttattttaa 240
tttgcatctg ggagaatgtc tgagcaagga gacctgaatc aggcaatagc agaggaagga 300
gggactgagc aggagacggc cactccagag aacggcattg ttaaatcaga aagtctggat 360
gaagaggaga aactggaact gcagaggcgg ctggaggctc agaatcaaga aagaagaaaa 420
tccaagtcag gagcaggaaa aggtaaactg actcgcagtc ttgctgtctg tgaggaatct 480
tctgccagac caggaggtga aagtcttcag ggtcagactc tctgaaaact gcaaatngga 540
agggatttca aaaggtttag gttaaaagtt aaattaaaag taggnacagt agtgctgaat 600
tttcctcaaa ggctctcttt tgataaggct gaaccaaata taatcccaag aatactctct 660
ccttccttgt tggagatgtc ttacctctca gctccccaaa atgcacttgc ctataagaaa 720
cacaattgct ggttcatata aacttaggaa atagtgaata aggtgcattt aactttggag 780
aaatactttt atggctttgg tggagatttc tcaatactgc caaagttgtc cagaaatcga 840
tctgagctga tggctgcttt tagttcatat tatcat 876

<210> 32
<211> 1521
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 829443CB1

<400> 32
caagctggcc ctgcacggct gcaagggagg ctcctgtgga caggccaggc aggtgggcct 60
caggaggtgc ctccaggcgg ccagtgggcc tgaggcccca gcaagggcta gggtccatct 120
ccagtcccag gacacagcag cggccaccat ggccacgcct gggctccagc agcatcagca 180
gcccccagga ccggggaggc acaggtggcc cccaccaccc ggaggagcag ctcctgcccc 240

tgtccggggg atgactgatt ctcctccgcc aggccaccca gaggagaagg ccaccccgcc 300
tggaggcaca ggccatgagg ggctctcagg aggtgctgct gatgtggctt ctggtgttgg 360
cagtgggcgg cacagagcac gcctaccggc ccggccgtag ggtgtgtgct gtccgggctc 420
acggggaccc tgtctccgag tcgttcgtgc agcgtgtgta ccagcccttc ctcaccacct 480
gcgacgggca ccgggcctgc agcacctacc gaaccatcta taggaccgcc taccgccgca 540
gccctgggct ggcccctgcc aggcctcgct acgcgtgctg ccccggctgg aagaggacca 600
gcgggcttcc tggggcctgt ggagcagcaa tatgccagcc gccatgccgg aacggaggga 660
gctgtgtcca gcctggccgc tgccgctgcc ctgcaggatg gcggggtgac acttgccagt 720
cagatgtgga tgaatgcagt gctaggaggg gcggctgtcc ccagcgctgc gtcaacaccg 780
ccggcagtta ctggtgccag tgttgggagg ggcacagcct gtctgcagac ggtacactct 840
gtgtgcccaa gggagggccc cccagggtgg cccccaaccc gacaggagtg gacagtgcaa 900
tgaaggaaga agtgcagagg ctgcagtcca gggtggacct gctggaggag aagctgcagc 960
tggtgctggc cccactgcac agcctggcct cgcaggcact ggagcatggg ctcccggacc 1020
ccggcagcct cctggtgcac tccttccagc agctcggccg catcgactcc ctgagcgagc 1080
agatttcctt cctggaggag cagctggggt cctgctcctg caagaaagac tcgtgactgc 1140
ccagcgcccc aggctggact gagcccctca cgccgccctg cagcccccat gcccctgccc 1200
aacatgctgg gggtccagaa gccacctcgg ggtgactgag cggaaggcca ggcagggcct 1260
tcctcctctt cctcctcccc ttcctcggga ggctccccag accctggcat gggatgggct 1320
gggatcttct ctgtgaatcc acccctggct acccccaccc tggctacccc aacggcatcc 1380
caaggccagg tgggccctca gctgagggaa ggtacgagct ccctgctgga gcctgggacc 1440
catggcacag gccaggcagc ccggaggctg ggtggggcct cagtgggggc tgctgcctga 1500
cccccagcac aataaaaatg t 1521

<210> 33
<211> 1349
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1470058CB1

<400> 33
aggctgggct gattcccttc tgtgctccca catgcttccg gaacagtgac aggtgagtga 60
aaaccaggta tcctgagccc catccctgat gtccctactc catccctgct atatccctag 120
cctgactttc cagcctcctg caggctcctg cctctaacca gcttctaccc cagccccatc 180
tctgcccctg ctagggactt cgcttttgtg gcaagtgaca aagatagctg tatgctcaag 240
tgccatgtgt ttcgctgtga tgtccctgcc aaggccattg ccagtgccct acatgggctt 300
tgtgcccaga tcttgtcaga gcgagtagag gtcagtggtg atgcctcttg ctgctcccca 360
gaccccatct ctcctgaaga cctgccacgg caagtggagc tgctggatgc ggtaagccaa 420
gctgctcaga agtacgaggc actgtatatg gggacactgc cagtcaccaa ggccatgggc 480
atggatgtgc tgaacgaggc cattggtacc ctcaccgcca gggggaccg gaatgcctgg 540
gtccccacca tgctcagtgt gtctgactct ctcatgactg cacatcccat tcaggcagag 600
gccagtacag aggaggagcc attgtggcag tgccctgtgc gccttgtgac atttattggt 660
gttggccgcg acccacacac ctttggcctc atcgctgacc tgggccgtca gagcttccag 720
tgcgcagcct tctggtgcca gccccatgca gggggactct ctgaagctgt gcaggctgcc 780
tgtatggttc agtaccagaa gtgtcttgtg gcctctgcag ctcgaggcaa ggcctggggt 840
gcccaggccc gtgcccgcct gcggctcaag cggaccagct ccatggattc cccaggaggt 900
cccctgcccc tcccccctgct caaaggaggg gttggcggtg caggggcaac ccctcgaaag 960
cggggtgtct tctcttttct tgatgccttc cggctgaaac cctctctgct ccatatgccc 1020
taaacttatc tgggaaggct ggggaagtag gctctgggtc catgcctaac tctgtaccgt 1080
tttattcctc aaggcctata gcctgtcact ccttgaagcc ttctctgcct gtccctccga 1140
tccttgtcca ccgtctattt attgcccaat ttattgttta tacggatgac tgggaggcac 1200
tgcaccacaa cgtaggaccc tggctcccct ttccttgggt ccttgtgttc cttgcccctg 1260

tccaaccctg gacagttggc tctacctcag taacacttta tagcaaaatc agtgcaaata 1320
aaaatccctc agtgacctca aaaaaaaaa 1349

<210> 34
<211> 1338
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1554947CB1

<400> 34
ggctgttgct gtggtttcct gagttgctgc tgctgcggcg gcggcagcgg cgtctgtgct 60
tgtggaggtg tcggcctctg ggcggatgtt gacattgtgt tgttgttatt gctgatggta 120
atggcggcgg cggtggcggc gacggtccag accccatccc ctctgtagcc ggagccgaga 180
cagccgacag cgaactccgc ggcctcggag ccggcggcag cggcgactcc cctcagcctc 240
cgccgcctcg cccgccggta ccccggcgcc aaccccggga gtcaggccct ttgggcaggg 300
gagctcggag gctcaggatg gcggatttcg acgaaatcta tgaggaagag gaggacgagg 360
agcgggccct ggaggagcag ctgctcaagt actcgccgga cccggtggtc gtccgcggct 420
ccggtcacgt caccgtattt ggactgagca acaaatttga atctgaattc ccttcttcat 480
taactggaaa agtagctcct gaagaattta aagccagcat caacagagtt aacagttgtc 540
ttaagaagaa ccttcctgtt aatgtacgtt ggctactttg tggctgcctt tgttgctgct 600
gcacattagg ttgcagtatg tggccagtta tttgcctcag taaaagaaca cgaagatcga 660
ttgagaagtt attagaatgg gaaaacaata ggttatacca caagctgtgc ttgcattgga 720
gactgagcaa aaggaaatgt gaaacgaata acatgatgga atatgtcatc ctcatagaat 780
ttttaccaaa gacaccgatt tttcgaccag attagcattt actttattta tagagacttt 840
ccaagtatgt tgtctttcca atggtgcctt gcttggtgct ctcctggtgg tgacataaca 900
ttggttctac agaatcgtgt ggtgtttttt ttgtttttgt tttttttttt tttttaaata 960
accgcatgtt ctaagtgtgc atttttgtca atctttgcaa cagttatttc atacagatgt 1020
ttaatactta agttattgtg ctcttttctg ttatgtattc tgattttcaa ggattacttt 1080
tttgtattat caaaaaaata catttgaact tagcataaaa agtggccagc ctttttttatt 1140
ttgtcaccaa ggtacacaca gtcctttatt tataaattcc ttaacagaga aaaacacctt 1200
tgtaaggctc aacttaccta ttccagcaag cacacttttt ctgtcatttt ttctttcttt 1260
tcaaatttga tattgtcatt attttaaaat agtaagtgtt ctttaatagt cttttgggac 1320
ctaacatacc ctttctca 1338

<210> 35
<211> 2120
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1690245CB1

<400> 35
tccctctcac caccctcttc tgcctatgag cggggaacaa aaaggccaga tgacagatat 60
gatacaccaa cttctaaaaa gaaagtacga attaaagacc gcaataaact ttctacagag 120
gaacgccgga agttgtttga gcaagaggtg gctcaacggg aggctcagaa acaacagcaa 180
cagatgcaga acctgggaat gacatcacca ctgccctatg actctcttgg ttataatgcc 240
ccgcatcatc cctttgctgg ttacccacca ggttatccca tgcaggccta tgtggatccc 300
agcaacccta atgctggaaa ggtgctcctg cccacaccca gcatggaccc agtgtgttct 360

```
cctgctcctt atgatcatgc tcagcccttg gtgggacatt ctacagaacc cctttctgcc 420
cctccaccag taccagtggt gccacatgtg gcagctcctg tggaagtttc cagttcccag 480
tatgtggccc agagtgatgg tgtagtacac caagactcca gcgttgctgt cttgccagtg 540
ccggcccccg gcccagttca gggacagaat tatagtgttt gggattcaaa ccaacagtct 600
gtcagtgtac agcagcagta ctctcctgca cagtctcaag caaccatata ttatcaagga 660
cagacatgtc caacagtcta tggtgtgaca tcaccttatt cacagacaac tccaccaatt 720
gtacagagtt atgcccagcc aagtcttcag tatatccagg ggcaacagat tttcacagct 780
catccacaag gagtggtggt acagccagcc gcagcagtga ctacaatagt tgcaccaggg 840
cagcctcagc ccttgcagcc atctgaaatg gttgtgacaa ataatctctt ggatctgccg 900
ccccccctctc ctcccaaacc aaaaaccatt gtcttacctc ccaactggaa gacagctcga 960
gatccagaag ggaagattta ttactaccat gtgatcacaa ggcagactca gtgggatcct 1020
cctacttggg aaagcccagg agatgatgcc agccttgagc atgaagctga gatggacctg 1080
ggaactccaa catatgatga aaaccccatg aaggcctcga aaaagcccaa gacagcagaa 1140
gcagacacct ccagtgaact agcaaagaaa agcaaagaag tattcagaaa agagatgtcc 1200
cagttcatcg tccagtgcct gaacccttac cggaaacctg actgcaaagt gggaagaatt 1260
accacaactg aagactttaa acatctggct cgcaagctga ctcacggtgt tatgaataag 1320
gagctgaagt actgtaagaa tcctgaggac ctggagtgca atgagaatgt gaaacacaaa 1380
accaaggagt acattaagaa gtacatgcag aagtttgggg ctgtttacaa acccaaagag 1440
gacactgaat tagagtgact gttgggccag ggtgggagga tgggtggtca ggtaagacag 1500
actctaggga gaggaaatcc tgtgggcctt tctgtcccac ccctgtcagc actgtgctac 1560
tgatgataca tcaccctggg gaattcaacc ctgcagatgt caactgaagg ccacaaaaat 1620
gaactccatc tacaagtgat tacctagttg tgagctgttg gcatgtggtt agaagccatc 1680
agaggtgcaa gggcttagaa aagaccctgg ccagacctga ctccactctt aaacctgggt 1740
cttctccttg gcggtgctgt cagcgcacag acccatgcgc atccccaccc acaacccttt 1800
accctgatga tctgtattat attttaatgt atatgtgaat atattgaaaa taatttgttt 1860
tttcctggtt tttgtttggt tttcgttttg cttttagcct ctacatgcta ggatcacagg 1920
aagactttgt aaggacagtt taagttctcc tgcaaggttt aatttgttat catgtaaata 1980
ttccaaagca ggctgccttg tggttttggc cagccttgtg ctatgttgat aagattgatt 2040
tactgcttaa aatcacttta ctttatccaa tttttactga actttttatg taaaaaaata 2100
aaatcaatta aagaaaaaaa                                             2120
```

<210> 36
<211> 642
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1878262CB1

<400> 36

```
ctcctgcact aggctctcag ccagggatga tgcgctgctg ccgccgccgc tgctgctgcc  60
ggcaaccacc ccatgccctg aggccgttgc tgttgctgcc cctcgtcctt ttacctcccc 120
tggcagcagc tgcagcgggc ccaaaccgat gtgacaccat ataccagggc ttcgccgagt 180
gtctcatccg cttgggggac agcatgggcc gcggaggcga gctggagacc atctgcaggt 240
cttggaatga cttccatgcc tgtgcctctc aggtcctgtc aggctgtccg gaggaggcag 300
ctgcagtgtg ggaatcacta cagcaagaag ctcgccaggc cccccgtccg aataacttgc 360
acactctgtg cggtgccccg gtgcatgttc gggagcgcgg cacaggctcc aaaaccaacc 420
aggagacgct gcgggctaca gcgcctgcac tccccatggc ccctgcgccc ccactgctgg 480
cggctgctct ggctctggcc tacctcctga ggcctctggc ctagcttgtt gggttgggta 540
gcagcgcccg tacctccagc cctgctctgg cggtggttgt ccaggctctg cagagcgcag 600
cagggctttt cattaaaggt atttatattt gtaaaaaaaa aa                    642
```

<210> 37
<211> 2536
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2253519CB1

<400> 37
gccgtggtgc tcaacgccct cacggtggac gcccacacag tcgtcagcca cccggacaag 60
tactgcttct actgccgggc gctgctgatg accgtggctg ggctgaagct gctgcgctca 120
gccttctgct gccccccaca gcagtacctg acgttggcct tcaccgtcct gctcttccac 180
tttgactacc cgcgcctctc ccagggcttt ctgcttgact acttcctcat gtccctgctt 240
tgcagcaagc tgtgggactt gctgtacaag ctgcgtttcg tgctgaccta catcgcgccc 300
tggcagatca cctggggctc ggctttccac gcttttgccc agccgtttgc cgtgccacac 360
tcggccatgc tgttcgtcca ggccctgctc tcggggctct tctccacgcc tctcaaccca 420
ctgctaggca gtgccgtctt catcatgtcc tacgctcggc ccctcaagtt ctgggagcgc 480
gactacaaca ctaaacgtgt ggatcattcc aacacccgcc tggtcacaca gctggacagg 540
aaccctggcg ctgatgacaa caacctcaac tccatcttct atgagcactt gacacgttcg 600
ctgcagcaca cactgtgtgg ggacctggtg ctgggccgct ggggcaacta tggccctggt 660
gactgcttcg tcctggcctc tgactacctc aacgccctgg tgcacctcat cgaggttggc 720
aatggcctcg tcaccttcca gctgcgtggc cttgagttcc ggggcactta ctgccagcag 780
cgcgaggtgg aggctatcac cgagggtgtg gaggaggacg agggctgttg ctgctgtgaa 840
cctggccacc tgccacgggt cctgtccttc aatgctgcct ttgggcagcg ctggctggct 900
tgggaggtaa cagccagcaa gtacgtgctg gagggctata gcattagtga caataatgct 960
gcctccatgc tgcaggtttt cgacctccgc aagatcctca tcacctacta tgtcaagagc 1020
atcatctact acgtgagccg ctcaccaaag ctggaggtgt ggctcagcca tgagggcatc 1080
acggcagccc tgaggcctgt gcgggtgccc ggctatgccg actcggatcc caccttctcg 1140
ctgagtgtgg atgaggacta tgacctccgc ctgtctggcc tctcgctgcc ctccttttgt 1200
gctgtgcacc tcgagtggat ccagtactgc gcctcccggc gcaccaggcc cgtggaccag 1260
gattggaact ccccgctggt cacgctgtgt tttggcctgt gtgtgctggg ccgccgggcc 1320
ctggggacag cctctcacag catgtctgca agcctggagc ccttcctcta cggcctgcac 1380
gccctgttca agggggattt tcgcatcacc tccccacgtg acgagtgggt ctttgccgac 1440
atggacctgc ttcaccgcgt tgtggcgcct ggggttcgca tggccctcaa gcttcaccag 1500
gaccacttca cgtccccaga tgaatatgag gagccagcag ccctatacga tgccattgcg 1560
gccaacgagg agcggctggt catctcacat gagggtgacc cagcatggcg cagcgccatc 1620
ctcagcaaca cgccctccct gctggcgctg cgccatgtcc tggatgatgc ctccgacgag 1680
tacaagatca tcatgctcaa ccggcgccac ctcagcttcc gagtcatcaa ggtgaaccgg 1740
gagtgcgtgc gcggcctgtg ggccgggcag cagcaggagc tggtgttcct gcgcaaccgc 1800
aaccccgagc gtggcagcat ccagaacgcc aagcaggcgc ttcgcaacat gatcaactcc 1860
tcctgtgacc agccgctggg ctaccccatc tacgtgtcgc ctctcaccac ctcgctggct 1920
ggcagccacc cccagctacg ggcactgtgg ggtggcccca tcagcctggg tgccattgcc 1980
cactggctcc tgcgcacctg ggagaggctt cacaagggct gtggcgccgg ctgcaatagt 2040
ggcgggaacg tggatgattc agactgtagt gggggcggtg gcctgacctc cctcagcaat 2100
aacccccccg tggcacaccc cacacctgag aacacggcag gcaatggtga ccaacccctc 2160
ccaccaggcc ctggctgggg gccgcggtcc tccctgagtg gctctggtga tggcggcccc 2220
ccacctctgc tgcagtggcc tccccctcgg ctccctggac cacccccctgc atcgcctatc 2280
cccacagagg gtccccggac ctcacggccc cctggcccgg gtctcctcag ttctgagggc 2340
cccagtggaa agtggagcct gggggccgg aaggggctgg gaggatctga cggggagcca 2400
gcctcaggga gccccaaagg aggtaccccc aaatctcagg taaggcacct gtgggagggt 2460
tgggtcccag aaggctaagg cctgctcacc cgccaacctc tccccctcc cccaggcgcc 2520
tctagacctc agcctc 2536

<210> 38
<211> 3957
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2888437CB1

<400> 38
ggaagctgca gagaaagtca gctggataaa ggacaaactt ctgcatattt attatcagaa 60
cagcattgac gacaaactgt tggtagagaa aatctttgct cagtatcttg tcccccacaa 120
cctggaaaca gaagagagaa tgaaatgctt atattactta tatgctagtt tggatccaaa 180
tgctgtaaaa gctctcaacg aaatgtggaa gtgtcagaac atgcttcgga tccatgtacg 240
cgaactattg gatttgcaca agcagcctac atcagaggct aactgttctg ccatgtttgg 300
aaaactgatg accatagcaa agaatttgcc tgaccccggg aaagcacaag attttgtgaa 360
gaaatttaac caggttctcg gcgatgatga gaaacttcgg tctcagttgg agttattaat 420
tagcccaacc tgttcttgca aacaagcaga tatttgtgtg agagaaatag cccggaaact 480
tgcaaatcct aagcaaccaa caaatccttt tctagagatg gtcaaatttc tgttggaaag 540
aatcgcacct gtgcacattg attcagaagc cataagtgca ctagtgaaat tgatgaataa 600
gtcaatagag gggacagcag atgatgaaga ggagggtgta agtccagata cagctatccg 660
ttcaggactt gaacttctta aggttctgtc ttttacacat cctacctcgt tccactctgc 720
agagacatat gagtccttgt tacagtgcct aagaatggag gatgacaagg tagcagaagc 780
tgctattcaa atttttagaa atacaggtca caaaatagaa acagaccttc cccagatacg 840
atcgacctta attcccattt tacatcaaaa agcaaagagg ggtactccac accaagcaaa 900
acaggctgtg cactgtatac acgccatatt cacaaataaa gaagtccagc ttgcacagat 960
ttttgagcca ctcagtagga gtctgaatgc tgatgtgcca gaacaactta taactccatt 1020
agtttcattg ggccacattt ctatgttagc accagatcag tttgcttccc caatgaaatc 1080
tgtagtagca aattttattg tgaaagatct gctaatgaat gacaggtcaa caggtgaaaa 1140
gaatggaaaa ctgtggtctc cagatgaaga ggtttcccct gaagtactag caaaggtaca 1200
ggcaattaaa cttctggtaa ggtggctgtt gggtatgaaa aacaaccagt ctaaatctgc 1260
caattcaacc cttcggttat tatcagcgat gttggttagt gagggtgacc tgacagagca 1320
aaagaggatc agtaaatctg atatgtctcg cttgcgatta gctgctggta gtgccataat 1380
gaagcttgct caggaacctt gttaccatga aattattacc ccagaacagt ttcagctctg 1440
tgcacttgtt attaatgatg agtgttacca agtaaggcag atatttgctc agaagctgca 1500
taaggcactt gtgaagttac tgctcccatt ggagtatatg gcgatctttg ccttgtgtgc 1560
caaagatcct gtgaaggaga gaagagcaca cgcacgacaa tgtttactga aaaatatcag 1620
tatacgcagg gaatacatta agcagaatcc tatggctact gagaaattat tatcactgtt 1680
gcctgaatat gtagttccat acatgattca cctgctagcc catgatccag attttacaag 1740
atcacaagat gttgatcagc ttcgtgatat caaagagtgc ctatggttca tgcttgaagt 1800
tttaatgaca aagaatgaaa acaatagcca tgcctttatg aagaagatgg cagagaacat 1860
caagttaacc agagatgccc agtctccaga tgaatccaag acaaatgaaa aactgtatac 1920
agtatgtgat gtggctctct gtgttataaa tagtaaaagt gctttgtgca atgcagattc 1980
accaaaggac ccagtcctcc caatgaaatt ttttacacaa cctgaaaagg acttctgtaa 2040
cgataagagt tatatttcag aagagacaag agtacttctg ttaacaggaa agccaaagcc 2100
tgctggagta ctaggtgcag taaataagcc tttatcagca acgggaagga aaccctatgt 2160
tagaagcact ggcactgaga ctggaagcaa tattaatgta aattcagagc tgaacccttc 2220
aaccggaaat cgatcaaggg aacagagttc agaggcagca gaaactggag ttagtgaaaa 2280
tgaagagaac cctgtgagga ttatttcagt cacacctgta aagaatattg acccagtaaa 2340
gaataaggaa attaattctg atcaggctac ccagggcaac atcagcagtg accgaggaaa 2400
gaaaagaaca gtaacagcag ctggtgcaga gaatatccaa caaaaaacag atgagaaagt 2460
agatgaatcg ggacctcccg cccccttccaa acccaggaga ggacgtcgac ccaagtctga 2520
atctcagggc aatgctacca aaaatgatga tctaaataaa cctattaaca agggaaggaa 2580
gagagctgca gtgggtcagg agagccctgg gggtttggaa gcaggtaatg ccaaagcacc 2640
caaactgcaa gatttagcca aaaaggcagc accagcagaa agacaaattg acttacaaag 2700

gtaaaaatgc atttgcaaag ggagaaaatg aaggccaaac agaagcaggc tccagcttct 2760
gcaaaaactt ggattcacaa atgtccctga acagaaaatg aagctcaacg cagaccgtgt 2820
gtgttctgaa gtgagcttca gaacacacac tctctgcctt gaaaactaaa gagactatta 2880
cttccttttc acatgaccac aagtcctctg atggaaatgt acagcagaaa ctcttgagag 2940
agaggctaaa agcaactctg ttctccccct tcccctagac ttttcttacg aaaagtcaat 3000
aattaagcaa attgcttaac acttggttcc agttcctgcc tatctggagt ttaaatgcgt 3060
aatacaccat taatttccac gctgcagttt ttattttaaa gaaagtaaca agatgtcttt 3120
acactgacac tgaaaattca tccattttag agccaggaat tcccatgtta cacaggaaaa 3180
aatagaagtc tactgaatta atttttaaa agaaaagaga tcagattaaa tatttctttg 3240
tttttccttt tggaaacttt tatgtataat tctttctgcc tgcctacttt tctgcaaaaa 3300
tgagatgtac agatttcggt tccctgctat gaaaagtgat gtggtagcaa ttttataaat 3360
gttgctttct gatttttatc agagtgagaa aattaaaatt attgatttgc aagtagtaaa 3420
cagttcatat tttgatttcc cctcatttta gtttaatata atttgcaata aatgtacata 3480
ttgttgtttg tttcataaag catatcactt taaaatggtt tttactcctg tgattatgtt 3540
ggaatatttg gaattttaaa ggagtaaaga ctgtccagca tttggtttta taatgtttgt 3600
caccagattt ttattaatgt aaaaaaaatc aatttttaaa aaatagttgg actttggcag 3660
cttttaagga aagttggagg tgttttagga ttgctatcaa ttttcagcat tgtgctattt 3720
ggaaataagt gtttgcttt tgtctgatgg tctgggctca tttttatgtt tattttagaa 3780
aactgttgca tcaatatatt atgtttcttg gcattgttca gcataggtaa tgtgtgcact 3840
ttatgtgtac acataatcat atttaagttt tttgcataaa ataaatgctt ctagatgtca 3900
tggcagtctt tttaatcttt ttatcatatg ctttcttgtg aatttttca tgttaaa 3957

<210> 39
<211> 3188
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3201753CB1

<400> 39
ggggagtcgg ctaagcaatg gagaagtgat gaggggcatt ttcatcaaac atgttctgga 60
agatagtcca gctggcaaaa atggaacctt gaaacctgga gatagaatcg tagaggtgga 120
tggaatggac ctcagagatg caagccatga acaagctgtg gaagccattc ggaaagcagg 180
caaccctgta gtctttatgg tacagagcat tataaacaga ccaagggcac ccagtcagtc 240
agagtcagag ccagagaagg ctccattgtg cagtgtgccc ccaccccctc cttcagcctt 300
tgccgaaatg ggtagtgatc acacacagtc atctgcaagc aaaatctcac aagatgtgga 360
caaagaggat gagtttggtt acagctggaa aaatatcaga gagcgttatg gaaccctaac 420
aggcgagctg catatgattg aactggagaa aggtcatagt ggtttgggcc taagtcttgc 480
tgggaacaaa gaccgatcca ggatgagtgt cttcatagtg ggattgatc caaatggagc 540
tgcaggaaaa gatggtcgat tgcaaattgc agatgagctt ctagagatca atggtcagat 600
tttatatgga agaagtcatc agaatgcctc atcaatcatt aaatgtgccc cttctaaagt 660
gaaaataatt tttatcagaa ataaagatgc agtgaatcag atggccgtat gtcctggaaa 720
tgcagtagaa cctttgcctt ctaactcaga aaatcttcaa aataaggaga cagagccaac 780
tgttactact tctgatgcag ctgtggacct cagttcattt aaaaatgtgc aacatctgga 840
gcttcccaag gatcaggggg gtttgggtat tgctatcagc gaagaagata cactcagtgg 900
agtcatcata aagagcttaa cagagcatgg ggtagcagcc acggatggac gactcaaagt 960
cggagatcag atactggctg tagatgatga aattgttgtt ggttacccta ttgaaaagtt 1020
tattagcctt ctgaagacag caaagatgac agtaaaactt accatccatg ctgagaatcc 1080
agattcccag gctgttcctt cagcagctgg tgcagccagt ggagaaaaaa agaacagctc 1140
ccagtctctg atggtcccac agtctggctc cccagaaccg gagtccatcc gaaatacaag 1200
cagatcatca acaccagcaa tttttgcttc tgatcctgca acctgcccca ttatccctgg 1260
ctgcgaaaca accatcgaga tttccaaagg gcgaacaggg ctgggcctga gcatcgttgg 1320

gggttcagac acgctgctgg gtgccattat tatccatgaa gtttatgaag aaggagcagc 1380
atgtaaagat ggaagactct gggctggaga tcagatctta gaggtgaatg gaattgactt 1440
gagaaaggcc acacatgatg aagcaatcaa tgtcctgaga cagacgccac agagagtgcg 1500
cctgacactc tacagagatg aggccccata caaagaggag gaagtgtgtg acaccctcac 1560
tattgagctg cagaagaagc cgggaaaagg cctaggatta agtattgttg gtaaaagaaa 1620
cgatactgga gtatttgtgt cagacattgt caaaggagga attgcagatg ccgatggaag 1680
actgatgcag ggagaccaga tattaatggt gaatggggaa gacgttcgta atgccaccca 1740
agaagcggtt gccgctttgc taaagtgttc cctaggcaca gtaaccttgg aagttggaag 1800
aatcaaagct ggtccattcc attcagagag gaggccatct caaagcagcc aggtgagtga 1860
aggcagcctg tcatctttca cttttccact ctctggatcc agtacatctg agtcactgga 1920
aagtagctca aagaagaatg cattggcatc tgaaatacag ggattaagaa cagtcgaaat 1980
gaaaaagggc cctactgact cactgggaat cagcatcgct ggaggagtag gcagcccact 2040
tggtgatgtg cctatattta ttgcaatgat gcacccaact ggagttgcag cacagaccca 2100
aaaactcaga gttggggata ggattgtcac catctgtggc acatccactg agggcatgac 2160
tcacacccaa gcagttaacc tactgaaaaa tgcatctggc tccattgaaa tgcaggtggt 2220
tgctggagga gacgtgagtg tggtcacagg tcatcagcag gagcctgcaa gttccagtct 2280
ttctttcact gggctgacgt caagcagtat atttcaggat gatttaggac ctcctcaatg 2340
taagtctatt acactagagc gaggaccaga tggcttaggc ttcagtatag ttggaggata 2400
tggcagccct catggagact tacccattta tgttaaaaca gtgtttgcaa agggagcagc 2460
ctctgaagac ggacgtctga aaggggcga tcagatcatt gctgtcaatg ggcagagtct 2520
agaaggagtc acccatgaag aagctgttgc catccttaaa cggacaaaag gcactgtcac 2580
tttgatggtt ctctcttgaa ttggctgcca gaattgaacc aacccaaccc ctagctcacc 2640
tcctactgta aagagaatgc actggtcctg acaattttta tgctgtgttc agccgggtct 2700
tcaaaactgt agggggggaaa taacacttaa gtttcttttt ctcatctaga aatgctttcc 2760
ttactgacaa cctaacatca tttttctttt cttcttgcat tttgtgaact taaagagaag 2820
gaatatttgt gtaggtgaat ctcgttttta tttgtggaga tatctaatgt tttgtagtca 2880
catgggcaag aattattaca tgctaagctg gttagtataa agaaagataa ttctaaagct 2940
aaccaaagaa aatggcttca gtaagttagg atgaaaaatg aaaatataaa ataaagaaga 3000
aaatctcggg gagtttaaaa aaaatgcctc aatttggcaa tctacctcct ctccccaccc 3060
caaactaaaa aaagaaaaaa aggttttcta atgaaaatct ttaaaaatac tgtcagtatt 3120
ttaaaatttt caacagtatt ataaaaacat tgcatctccc cacctctaat atgcatatat 3180
atttttcc 3188

<210> 40
<211> 3551
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3800639CB1

<400> 40
tgtctgaggc ggcgctgtgt gtgtgaagcg tacctagggc gggaggcgac atggagacag 60
gggcggccga gctgtatgac caggcccttt tgggcatcct gcagcacgtg ggcaacgtcc 120
aggatttcct gcgcgttctc tttggcttcc tctaccgcaa gacagacttc tatcgcttgc 180
tgcgccaccc atcggaccgc atgggcttcc cgccccgggc cgcgcaggcc ttggtgctgc 240
aggtattcaa aacctttgac cacatggccc gtcaggatga tgagaagaga aggcaggaac 300
ttgaagagaa aatcagaaga aaggaagagg aagaggccaa gactgtgtca gctgctgcag 360
ctgagaagga gccagtccca gttccagtcc aggaaataga gattgactcc accacagaat 420
tggatgggca tcaggaagta gagaaagtgc agcctccagg ccctgtgaag gaaatggccc 480
atggttcaca ggaggcagaa gctccaggag cagttgctgg tgctgctgaa gtccctaggg 540
aaccaccaat tcttcccagg attcaggagc agttccagaa aaatcccgac agttacaatg 600
gtgctgtccg agagaactac acctggtcac aggactatac tgacctggag gtcagggtgc 660

cagtacccaa gcacgtggtg aagggaaagc aggtctcagt ggcccttagc agcagctcca 720
ttcgtgtggc catgctggag gaaaatgggg agcgcgtcct catggaaggg aagctcaccc 780
acaagatcaa cactgagagt tctctctgga gtctcgagcc cgggaagtgc gttttggtga 840
acctgagcaa ggtgggcgag tattggtgga acgccatcct ggagggagaa gagcccatcg 900
acattgacaa gatcaacaag gagcgctcca tggccaccgt ggatgaggag gaacaggcgg 960
tgttggacag gcttaccttt gactaccacc agaagctgca gggcaagcca cagagccatg 1020
agctgaaagt ccatgagatg ctgaagaagg ggtgggatgc tgaaggttct cccttccgag 1080
gccagcgatt cgaccctgcc atgttcaaca tctccccggg ggctgtgcag ttttaatgac 1140
cagaaggaaa ggaaaccctc gccggtgggg aggcagagcc ttatcctcgg ctgcccttct 1200
tggctccctg cattccaggg acttgctcgt cttgtttacc cctagccatc ctttctttca 1260
agggtgaacc aggccttcca ccctgacctt gcatctccag actgttccag agaaggtgcg 1320
gggccagctg ctatgtggtg gccgctgtgg ctgacactga gtgaaggtgt ttgaaatgca 1380
ggagaggata tcccagcaaa ttgggatcac atgcttttgt ctccacagca accagccact 1440
gcaggcagca tgtctttcct cccctgctct ctgcttgctg ttgttttgac gctattctgc 1500
ttgcatgtct tctggttggg atgtggagtt gttgctggac tctcaggcga actgaagtca 1560
ttgaagtgtg tgaagctctg tgcttgcatg agggcaagca aggaatggct gtgcctgagg 1620
ctgctctggg aaactccttg ccccttgacc tcttttgaga gcattcacgt ggtcttcttg 1680
ctcatcccct tataaatgtg ctttgcctgc ctcagcctca tggtcagagc agtggagact 1740
ggagccctgt ttgcacgttc tagttgttcg gagaaagcct aggttctggg ctcaggtcca 1800
gatgcagcgg ggattctgtt ctctgactgt ggcgaccttg ctttggttct tgttgaagtg 1860
aaccaagccc ggccaccacg catggcatgc tgtgcttggc tccccataag acgtcctctt 1920
tgggtgcacg gtgtcaaagt gtgggcagga gtggagagct ggtgccctca ggaggagacc 1980
acagcatgtc catcagctca gcagagctcg acagccacaa gtcctgagaa gctttgacct 2040
tgaagggctt ctgggagagg aggaatttct gcatggggcg tgaaggcaca ctgtcccacc 2100
acaactgaac cagaagagag tgaagactcc cctcttccca tcctctgtgc caggtgccag 2160
actgtgctcc ttggaactta tggcccaatc ttacctgttc tccagggact ggtcactgcc 2220
tcaggacccc caagcctatg ccctgagcca tggctgctga ctgactccag ccaaggtgca 2280
aagacgagat tatgagacag gtcctcaggc ctgtgttcca agtactcaca ggggctctgg 2340
gtgcccatcg ccgggagtat ggttcagctg ccaccggcac tgtccatttg cctgtctgtc 2400
aagctcagag catggataag ccacacagca gggcagtgca ccctggcacc atgcacggcc 2460
agcaagaatc aaggcccgca gatgctaaga gggcctattg tcaggggaag gtccccgctc 2520
ctgcacactc tctatggata cttgggttgt gggggctctc ttggagagta agtttgtggt 2580
ttgtttctgg tttacagtgg tggctgacac cccttgtaag aaagcattcc tgggaagtct 2640
tctgtgggtc caaacatgtt gctccgatca tcacaggaga gcaaaaggcc ctagataccc 2700
cctttggaat gtgagagtct tgttgtctga tatttgccac tgagctggtg aagcccctct 2760
aaagagatct cgaccctggg gagcagaatt cttgtcatct atgaggggtc ctgagaaaga 2820
cttgtcattt tttttcctgg agttcttccc attgaggtcc taggatttgc acaccactgt 2880
cccacaagag ctttcctgcc taatgaaagg aggtcttgtg gtgtgtgtct cctctcttct 2940
ctatagttcc cgagttggcc cccattgcag cccccacccct gtgggtagtc ttccagaagt 3000
gatgcagtgg tgtgagatgc cctacacctt gttatttggg agactttgag agtcattcac 3060
ttccatggtg actagtgttt gttttgcctg attttatatt ctgtgttgca tttctcccca 3120
ctccctgccc tgctttaata aacagcaaac caatatctag gaagaatgac tgagggatag 3180
tattgggtat tggccccatg gcaggaacag ccacttgcat ctggtcccgg tgccacactg 3240
cggtgcttgg tgtggttgtg gagcctgtcc ctgcgcgcct tgctcccgtt gagccacgct 3300
gtctggtggg tgattctctg ccctgagcca ccaccctgga ctggcccagt ctccagagct 3360
ggcacaccct gcctgttttc tctttttaga cacaacagcc gcagtttggc cagccactaa 3420
gtcccaccag ctgaggtccg aggaaagcgg ggtgactcat ttcccttgtc cagggcccga 3480
ggagagtgag gtgtccagcc tgcaaagcta ttccagctcc ttggtgttgg tttgcaataa 3540
attggtattt a 3551

<210> 41
<211> 2308
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 533825CB1

<400> 41
cgccaaccgt ggtggctcct tgcgttccta catcctctca tctgagaatc agagagcata 60
atcttcttac gggcccgtga tttattaacg tggcttaatc tgaaggttct cagtcaaatt 120
ctttgtgatc tactgattgt gggggcatgg caaggtttgc ttaaaggagc ttggctggtt 180
tgggcccttg tagctgacag aaggtggcca gggagaaggc agcacactgc tcggagaatg 240
aaggcgcttc tgttgctggt cttgccttgg ctcagtcctg ctaactacat tgacaatgtg 300
ggcaacctgc acttcctgta ttcagaactc tgtaaaggtg cctcccacta cggcctgacc 360
aaagatagga agaggcgctc acaagatggc tgtccagacg gctgtgcgag cctcacagcc 420
acggctccct ccccagaggt ttctgcagct gccaccatct ccttaatgac agacgagcct 480
ggcctagaca accctgccta cgtgtcctcg gcagaggacg ggcagccagc aatcagccca 540
gtggactctg gccggagcaa ccgaactagg gcacggccct ttgagagatc cactattaga 600
agcagatcat ttaaaaaaat aaatcgagct ttgagtgttc ttcgaaggac aaagagcggg 660
agtgcagttg ccaaccatgc cgaccagggc agggaaaatt ctgaaaacat cactgcccct 720
gaagtctttc caaggttgta ccacctgatt ccagatggtg aaattaccag catcaagatc 780
aatcgagtag atcccagtga aagcctctct attaggctgg tgggaggtag cgaaacccca 840
ctggtccata tcattatcca acacatttat cgtgatgggg tgatcgccag agacggccgg 900
ctactgccag gagacatcat tctaaaggtc aacgggatgg acatcagcaa tgtccctcac 960
aactacgctg tgcgtctcct gcggcagccc tgccaggtgc tgtggctgac tgtgatgcgt 1020
gaacagaagt tccgcagcag gaacaatgga caggccccgg atgcctacag accccgagat 1080
gacagctttc atgtgattct caacaaaagt agccccgagg agcagcttgg aataaaactg 1140
gtgcgcaagg tggatgagcc tggggtttc atcttcaatg tgctggatgg cggtgtggca 1200
tatcgacatg gtcagcttga ggagaatgac cgtgtgttag ccatcaatgg acatgatctt 1260
cgatatggca gcccagaaag tgcggctcat ctgattcagg ccagtgaaag acgtgttcac 1320
ctcgtcgtgt cccgccaggt tcggcagcgg agccctgaca tctttcagga agccggctgg 1380
aacagcaatg gcagctggtc cccagggcca ggggagagga gcaacactcc caagcccctc 1440
catcctacaa ttacttgtca tgagaaggtg gtaaatatcc aaaaggaccc cggtgaatct 1500
ctcggcatgg ccgtcgcagg gggagcatca catagagaat gggatttgcc tatctatgtc 1560
atcagtgttg agcccggagg agtcataagc agagatggaa gaataaaaac aggtgacatt 1620
ttgttgaatg tggatggggt cgaactgaca gaggtcagcc ggagtgaggc agtggcatta 1680
ttgaaaagaa catcatcctc gatagtactc aaagctttgg aagtcaaaga gtatgagccc 1740
caggaagact gcagcagccc agcagccctg gactccaacc acaacatggc cccacccagt 1800
gactggtccc catcctgggt catgtggctg gaattaccac ggtgcttgta taactgtaaa 1860
gatattgtat tacgaagaaa cacagctgga agtctgggct tctgcattgt aggaggttat 1920
gaagaataca atggaaacaa acctttttc atcaaatcca ttgttgaagg aacaccagca 1980
tacaatgatg gaagaattag atgtggtgat attcttcttg ctgtcaatgg tagaagtaca 2040
tcaggaatga tacatgcttg cttggcaaga ctgctgaaag aacttaaagg aagaattact 2100
ctaactattg tttcttggcc tggcactttt ttatagaatc aatgatgggt cagaggaaaa 2160
cagaaaaatc acaaatatag gctaagaagt tgaaacacta tatttatctt gtccagtttt 2220
atatttaaag aaagaataca ttgtaaaaat gtcaggaaaa gtatgatcat ctaatgaaag 2280
ccagttacac ctcagaaaat atgattcc 2308

<210> 42
<211> 1881
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1311833CB1

<400> 42
ccacactcgc ccctgcaagc aagcttgaaa tgggcttctg tctgagcggc accgggatga 60
ccaatgggtc tgtgttttgg accggcgaac caatggcgcg acagggaagg cgggcccgag 120
atgggttagg gtgtccttgc agggtgtctg agctaaactt caccaaataa tagctgtttg 180
tattttggct gctgcaggag ccattttaga aataaatatc ttccttcaat agatgaaaat 240
gaaaatacag aaaaaagaga agcagttgtc aaatttaaaa gttttgaatc actccccaat 300
gtctgatgcc tctgtcaatt ttgactacaa atctccatcc ccatttgact gcagcactga 360
tcaagaagag aaaattgaag atgttgctag tcactgtctg cctcagaagg acctgtatac 420
tgctgaagag gaagctgcta cccttttccc taggaaaatg acatcccata atgggatgga 480
ggacagtgga ggaggaggta ctggagtgaa gaagaaacgg aagaaaaagg agccaggaga 540
ccaagagggt gcagcaaagg gaagcaagga cagagagccc aagccaaaga ggaaacgaga 600
accgaaagag ccaaaggaac ccagaaaggc caaggagccg aagaaggcca aggagcacaa 660
ggagccgaag caaaaagatg gggcaaagaa ggcacggaag ccccgggagg cctcgggcac 720
caaggaggcc aaagagaaga ggagctgcac tgactctgca gccaggacga agtccaggaa 780
ggccagcaag gagcaaggac caaccccagt ggagaaaaag aagaaaggaa aaaggaaaag 840
tgaaactaca gtggagagtt tagagctgga tcagggcctg acgaacccat ctctgcggag 900
tcctgaggag tccactgagt ctacagacag ccagaaacga cgctcgggaa ggcaagtaaa 960
gcgcagaaaa tacaatgagg acctggactt caaagtggtg gatgatgatg gggaaacaat 1020
tgctgttctt ggagctggtc gaacatctgc actctcagcc tctacactgg cctggcaggc 1080
ggaggagcct ccagaagatg atgcaaacat cattgagaag atcctggcat ctaagactgt 1140
ccaggaggtt cacccaggag aacctccgtt cgacttggag ctgttctacg ttaagtatag 1200
aaatttttcc tacttacatt gtaaatgggc cacaatggaa gagctcgaaa aggatcctcg 1260
catcgcacag aagatcaagc gatttaggaa taaacaagcc cagatgaagc acatttttac 1320
ggaggtgaag caatatttac tgactcattt gactgctgct tttcttgcag cagtaaatac 1380
tgtgtttacg tttctaagtc caagttaaat tcctatacag atctgttgct caagtcaaat 1440
tgataaatgt gtgctgattt gcaacaaaac tagtctagat tcaagtcaaa gaaaactatc 1500
caccaaagtt aaggatcaaa ccttataaac tgacaactca taacggcgtg catcttggtc 1560
tgaatactaa taaatagcta gagtgagagg aatacgcaca caattctttt gtccctgctc 1620
cctcaaccca ctgcaaccca aaaaggacat ttcagaaggt cctaaagttt gaatgtaaaa 1680
aacaaaacaa atctaaaagt gagagaaagt aaacaaataa ataattttat ttatgtaaac 1740
ttgggaattt ctaagtattg tacctatagc aagaggaagg actgatggat tttaatatta 1800
aacaaatcta taatgctata tatcagaaaa tgccacaaac cttaacctta atgtgttatt 1860
cttggaataa agactaaacc a 1881

<210> 43
<211> 1974
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1342819CB1

<400> 43
ccagaagcca gcagtggggt tgcacacgcg cctcttcacg aggtggaaac aagatggagg 60
attcggcctc ggcctcgctg tcttctgcag ccgctactgg aacctccacc tcgactccag 120
cggccccgac agcacggaag cagctggata aagaacaggt tagaaaggca gtggacgctc 180
tcttgacgca ttgcaagtcc aggaaaaaca attatgggtt gcttttgaat gagaatgaaa 240
gtttattttt aatggtggta ttatggaaaa ttccaagtaa agaactgagg gtcagattga 300
ccttgcctca tagtattcga tcagattcag aagatatctg tttatttacg aaggatgaac 360
ccaattcaac tcctgaaaag acagaacagt tttatagaaa gcttttaaac aagcatggaa 420
ttaaaaccgt ttctcagatt atctccctcc aaactctaaa gaaggaatat aaatcctatg 480
aagccaagct ccgccttctg agcagttttg atttcttcct tactgatgcc agaattaggc 540
ggctcttacc ctcactcatt gggagacatt tctatcaaag aaagaaagtt ccagtatctg 600

taaaccttct gtccaagaat ttatcaagag agatcaatga ctgtataggt ggaacggtct 660
taaacatttc taaaagtggt tcttgcagtg ctatacgtat tggtcacgtt ggaatgcaaa 720
ttgagcacat cattgaaaac attgttgctg tcaccaaagg actttcagaa aaattgccag 780
agaagtggga gagcgtgaaa ctcctgtttg tgaaaactga gaaatcggct gcacttccca 840
tctttttcctc gtttgtcagc aattgggatg aagccaccaa aagatctttg cttaataaga 900
agaaaaaaga ggcaaggaga aaacgaagag aaagaaattt tgaaaaacaa aaggagagga 960
agaagaagag gcagcaggct aggaagactg catcagttct tagtaaagat gatgtggcac 1020
ctgaaagtgg tgatactaca gtgaagaaac ctgaatcaaa gaaggaacag accccagagc 1080
atgggaagaa aaaacgtggc agaggaaaag cccaagttaa agcaacaaat gaatccgaag 1140
acgaaatccc acagctggta ccaataggaa agaagactcc agctaatgaa aaagtagaga 1200
ttcaaaaaca tgccacagga aagaagtctc cagcaaagag tcctaatccc agcacacctc 1260
gtgggaagaa aagaaaggct ttgccagcat ctgagacccc aaaagctgca gagtctgaga 1320
ccccagggaa aagcccagag aagaagccaa aaatcaaaga agaggcagtg aaggaaaaaa 1380
gtccttcgct ggggaaaaaa gatgcgagac agactccaaa aaagccagag gccaagtttt 1440
tcaccactcc tagtaaatct gtgagaaaag cttcccacac ccccaaaaaa tggcccaaaa 1500
aacccaaagt accccagtcg acctaaagtc agtgattcaa ctggaaggaa acctcaatgc 1560
tgcctccaga gctttttgga aatactcaga tcctggccgc ctttgtaacc ttctctaaac 1620
gtcaggcctg gacttaaaag atttttaaa acctccataa gtagtccagg ggcggtggct 1680
cacgcctgta atcccagcac tttgggaggc cgaggcaggc ggatcacaag gtcaacgaga 1740
tcgagaccat cctggccaac atggtgaaac cctgtctgta ccaaaaatac aaaaattaat 1800
tgggcatggt ggtggacacc tgtaatccca gctactaggg aggctgaggc aggagaattg 1860
cttgaacctg ggaggcggag gttgcagtga gccactgcac tccagcctga tgacagagca 1920
agactcagtc tcaaaaataa ataaaaataa taaaacctca aaaaaaaaaa aaaa 1974

<210> 44
<211> 1061
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1871288CB1

<400> 44
tcctccttgg gttcgggtga aagcgcttgg gggttcagtg ggccatgatc cccgagctgc 60
tggagaactg aaggcggacg gtctcctgcg aaaccaggca atggcggagc tggagtttgt 120
tcagatcatc atcatcgtgg tggtgatgat ggtgatggtg gtggtgatca cgtgcctgct 180
gagccactac aagctgtctg cacggtcctt catcagccgg cacagccagg ggcggaggag 240
agaagatgcc ctgtcctcag aaggatgcct gtggccctcg gagagcacag tgtcaggcaa 300
cggaatccca gagccgcagg tctacgcccc gcctcggccc accgaccgcc tggccgtgcc 360
gcccttcgcc cagcgggagc gcttccaccg cttccagccc acctatccgt acctgcagca 420
cgagatcgac ctgccgccca ccatctcgct gtcagacggg gaggagcccc caccctacca 480
gggcccctgc accctccagc ttcgggaccc cgagcagcag ctggaactga accgggagtc 540
ggtgcgcgca cccccaaaca gaaccatctt cgacagtgac ctgatggata gtgccaggct 600
gggcggcccc tgccccccca gcagtaactc gggcatcagc gccacgtgct acggcagcgg 660
cgggcgcatg gaggggccgc cgcccaccta cagcgaggtc atcggccact acccggggtc 720
ctccttccag caccagcaga gcagtgggcc gccctccttg ctggagggga cccggctcca 780
ccacacacac atcgcgcccc tagagagcgc agccatctgg agcaaagaga aggataaaca 840
gaaaggacac cctctctagg gtccccaggg gggccgggct ggggctgcgt aggtgaaaag 900
gcagaacact ccgcgcttct tagaagagga gtgagaggaa ggcgggggc gcagcaacgc 960
atcgtgtggc cctcccctcc cacctccctg tgtataaata tttacatgtg atgtctggtc 1020
tgaatgcaca agctaagaga gcttgcaaaa aaaaaaaaaa a 1061

<210> 45
<211> 505
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2587338CB1

<400> 45
ggagcctgga ggagcccacg cagtctgttc ccggcacccg gtgcgtgtga agggacttga 60
gggcagcgag atggaatcag caagagaaaa catcgacctt caacctggaa gctccgaccc 120
caggagccag cccatcaacc tgaaccatta cgccaccaag aagagcgtgg cggagagcat 180
gctggacgtg gccctgttca tgtccaacgc catgcggctg aaggcggtgc tggagcaggg 240
accatcctct cactactaca ccaccctggt caccctcatc agcctctctc tgctcctgca 300
ggtggtcatc ggtgtcctgc tcgtggtcat tgcacggctg aacctgaatg aggtagaaaa 360
gcagtggcga ctcaaccagc tcaacaacgg cagccacatc ttggtcttct tcactgtggt 420
catcaatggt ttcattacag gcttcggggc acataaaaca agggtcctgg cctgccagga 480
ctccaggaat cctctctgaa tggag 505

<210> 46
<211> 1099
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2821211CB1

<400> 46
gaaatgcttt gtcttcagcc tctccaggca ccatctccct tcctgtggga gcagagagct 60
tagcctggag cacctttcct tcaagccagc aacacagagc actaggttca attccctgaa 120
ggtggccact ttaagagaga aatctgaaaa ccccatttgc tttcttttct cccatattgg 180
catggatttc tgtcttctct aacaccttgt gaccttctct atatcatgct ttaaagtgta 240
ataatatgat tttttaaaag aaatttatta cttgttgcaa aggtcttttt aaaccagttt 300
agatttcaag aaaaaataaa tggaaatcat cgaaaattca tttcacatta atggtctaaa 360
aataaaccaa aggacattat gtgtgcatgt gtgtataagt gcacacagaa atatatatac 420
atatgtagac tatatacatg tgtgtatata tgtgtatata tacatacact tgtataaatg 480
tatatacaca tatacctata atgtgtgtat gtgtatttat tgaagaaaca gataccatac 540
tcatttctaa aagaatattc agagaatatc aagatgattc tggctgaaaa aggccagtgg 600
aaattcaggt gaaaatgttc atcaattccc attgcatcac ctctgtaatt tttcagctct 660
ctgtataaac attaaatgtc ttatatagca gcaaaaatat aaaatagttg tccatatttt 720
cacaggtgtg gtgtaattta taaaattaga aagcaactta tcagctactt aagagaaatg 780
gcaagttttg atatgagtat acaatatata aaaatatata tagtgctata tatataaata 840
tttggtctct atttcatttt ttgcatcagt attaatacta aaatatgtct cgctagtgat 900
gtttttatga tatccctgat cctaactgaa gagacagtta tttatagtca tttattttaa 960
aaaatgaaaa taagtgaata ataattaggt taacattgtt gctccctgtg acaaaatttt 1020
ataagcaaat ttcaaaagac atgttgtaaa ttaggaggct caacaataaa acattatgct 1080
ccagaaaaaa aaaaaaaaa 1099

<210> 47
<211> 1727
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2824832CB1

<400> 47
gttgatccgc ggctgcgctc catgttccag tttcatgcag gctcttggga aagctggtgc 60
tgctgctgcc tgattcccgc cgacagacct tgggaccggg gccaacactg gcagctggag 120
atggcggaca cgagatccgt gcacgagact aggtttgagg cggccgtgaa ggtgatccag 180
agtttgccga agaatggttc attccagcca acaaatgaaa tgatgcttaa attttatagc 240
ttctataagc aggcaactga aggaccctgt aaactttcaa ggcctggatt ttgggatcct 300
attggaagat ataaatggga tgcttggagt tcactgggtg atatgaccaa agaggaagcc 360
atgattgcat atgttgaaga aatgaaaaag attattgaaa ctatgccaat gactgagaaa 420
gttgaagaat tgctgcgtgt cataggtcca ttttatgaaa ttgtcgagga caaaaagagt 480
ggcaggagtt ctgatataac ctcagatctt ggtaatgttc tcacttctac tccgaacgcc 540
aaaaccgtta atggtaaagc tgaaagcagt gacagtggag ccgagtctga ggaagaagag 600
gcccaagaag aagtgaaagg agcagaacaa agtgataatg atataaatga tgatcatgtt 660
gaagatgtta caggaattca gcatttgaca agcgattcag acagtgaagt ttactgtgat 720
tctatggaac aatttggaca agaagagtct ttagacagct ttacgtccaa caatggacca 780
tttcagtatt acttgggtgg tcattccagt caacccatgg aaaattctgg atttcgtgaa 840
gatattcaag tacctcctgg aaatggcaac attgggaata tgcaggtggt tgcagttgaa 900
ggaaaaggtg aagtcaagca tggaggagaa gatggcagga ataacagcgg agcaccacac 960
cgggagaagc gaggcggaga aactgacgaa ttctctaatg ttagaagagg aagaggacat 1020
aggatgcaac acttgagcga aggaaccaag ggccggcagg tgggaagtgg aggtgatggg 1080
gagcgctggg gctccgacag agggtcccga ggcagcctca atgagcagat cgccctcgtg 1140
ctgatgagac tgcaggagga catgcagaat gtccttcaga gactgcagaa actggaaacg 1200
ctgactgctt tgcaggcaaa atcatcaaca tcaacattgc agactgctcc tcagcccacc 1260
tcacagagac catcttggtg gcccttcgag atgtctcctg gtgtgctaac gtttgccatc 1320
atatggcctt ttattgcaca gtggttggtg tatttatact atcaaagaag gagaagaaaa 1380
ctgaactgag gaaaatggtg ttttcctcaa gaagactact ggaactggat gacctcagaa 1440
tgaactggat tgtggtgttc acaagaaaat cttagtttgt gatgattaca ttgctttttg 1500
ttgtccagta gtttagtttg tgtacatata tacacatata tattttgcac tacacaaacg 1560
ataacatttt aaggactaat attgctgata cttgaataat caatctctac taggttataa 1620
gtagtataca cagatttacc ctgcccttga acttgaagga cattaaatta ttaatgatca 1680
tttggtaaca tgtttacctg attatcttcc atagagtaac ataaggg 1727

<210> 48
<211> 951
<212> DNA
<213> Homo sapiens

<220>
<221> misc_Feature
<223> Incyte ID No.: 3070147CB1

<400> 48
ggggactgga gcatgggacg gcgcgcctga aggagcagga aggggaagga ggcctgggac 60
cccgaaaaga gaaggggaga gcgaggggac gagagcggag gaggaagatg caactgactc 120
gctgctgctt cgtgttcctg gtgcagggta gcctctatct ggtcatctgt ggccaggatg 180
atggtcctcc cggctcagag gaccctgagc gtgatgacca cgagggccag ccccggcccc 240
gggtgcctcg gaagcggggc cacatctcac ctaagtcccg ccccatggcc aattccactc 300
tcctagggct gctggccccg actggggagg cttggggcat tcttgggcag ccccccaacc 360
gcccgaacca cagcccccca ccctcagcca aggtgaagaa aatctttggc tggggcgact 420
tctactccaa catcaagacg gtggccctga acctgctcgt cacagggaag attgtggacc 480
atggcaatgg gaccttcagc gtccacttcc aacacaatgc cacaggccag ggaaacatct 540

ccatcagcct cgtgcccccc agtaaagctg tagagttcca ccaggaacag cagatcttca 600
tcgaagccaa ggcctccaaa atcttcaact gccggatgga gtgggagaag gtagaacggg 660
gccgccggac ctcgctttgc acccacgacc cagccaagat ctgctcccga gaccacgctc 720
agagctcagc cacctggagc tgctcccagc ccttcaaagt cgtctgtgtc tacatcgcct 780
tctacagcac ggactatcgg ctggtccaga aggtgtgccc agattacaac taccatagtg 840
ataccccсta ctacccatct gggtgacccg gggcaggcca cagaggccag gccagggctg 900
gaaggacagg cctgcccatg caggagacca tctggacacc gggcagggaa g 951

<210> 49
<211> 1624
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3271841CB1

<400> 49
ctcctccgcg ttcgcagccg ccgtcatccc gcggaggagc gcgcacgccc ggggaggccg 60
gaggacgcgc ccatagaatg cccaggggcg acaaactgtc ctgagcccct ctggtgcagc 120
cacctgcctg tcccatacgc cccgcccacc atggagtcca gagggaagtc agccagcagc 180
cccaagcccg acaccaaggt gccccaggtc accaccgagg ccaaggtacc cccggcagcc 240
gatgggaaag ccccсttgac caagccctcg aagaaggagg ccccggccga gaagcagcag 300
ccgccagcag cccccaccac ggcacctgcc aagaagacct cggccaaggc cgaccctgcc 360
cttctcaaca accacagcaa cctgaagcca gcccccacgg tccccagcag tcccgatgca 420
accccggagc ccaagggtcc tggggacggg gcggaggaag atgaggctgc cagtgggggg 480
cctgggggcc gaggtccctg gtcctgtgag aacttcaacc ccctgctggt ggctggggt 540
gtggccgtgg cagccatagc cctgattctc ggtgtggcct tcctggtccg gaaaaaataa 600
tacctggggg ccaggcgggg ggcacggagc cacttcctgt acagacccga ggaagccagt 660
gcatgcagag ttcaccctta cctattcgta cacacgcaca ttcattacac acctacatat 720
gcccccaaca cacgcgcacg gtgaagagga cgcccgagcc cacccctgct gacccaggac 780
ttccccaacc tccagggcag aaagagccca ggctccgggg tccacagcac aggatgtggg 840
gcgagggcac agctggggaa cggcaagaaa ggaatggacc ctgtgtgtgg cccccccacc 900
cctggcggct gggtgatcct gggcccccag ggctggtctg agtgcaggtg ggggtgccgg 960
gatgggttgg gcctgggccc ggccctcgtg gggacattaa agggcgcggt ggctccactc 1020
gccccatttc tgctctttgc gtaccccccg ggggtctgcc tgggcgaatg cagtgggacg 1080
gacctggggt gggcactcac ctcagcagag cctcaatgcc acctccccac accacccaag 1140
cctagctggg tgggaagacg gagccctgaa gcctgcgtgt ggagggggtg ctgcccaagg 1200
gtcgccttcc ccagcttgtg ctccagggga gccccgacgc cggcggggag gcagggatgt 1260
cgagggcccc tcacagctgt gcccgccctg ccccgctaga ctgcccctcc agcttgcacc 1320
cagcaccacc tgagtctaac cagcgtataa tgcaataaca ggtagagtag aactgctttt 1380
ggcggcgacc gtcacacact gtcctcagcc cactgcacct ggggaagccc ctcgtgcagc 1440
ccctccttag cacatccagt cccagcgtgg cccttggcgc gcagcagcag gcgaactccc 1500
agaaggtggt ggtgggcact gtgagaacgt ggcctgccct cactggccag tggccacagg 1560
aaccccgggc ctgagtggag gctgagtctg aaataaactc tgtcgtctga aaaaaaaaaa 1620
aaaa 1624

<210> 50
<211> 2080
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3537827CB1

<400> 50
gtcaaggtca ccttccctct ccagtagcag caggactgag tgattccatc tatttggagt 60
cgcggaagcg gtgaattgga ggggaggagc tgacccagat gaggaaactg agacccagaa 120
aggtggaagc acttgtctaa ggtcacgcct ccaggaagca gtgtgtccac gactccagtc 180
caagtggtca ggctccagag cccacagtcc caggggtcca tgatgccgag ctgcaatcgt 240
tcctgcagct gcagccgcgg ccccagcgtg gaggatggca agtggtatgg ggtccgctcc 300
tacctgcacc tcttctatga ggactgtgca ggcactgctc tcagcgacga ccctgaggga 360
cctccggtcc tgtgcccccg ccggccctgg ccctcactgt gttggaagat cagcctgtcc 420
tcggggaccc tgcttctgct gctgggtgtg gcggctctga ccactggcta tgcagtgccc 480
cccaagctgg agggcatcgg tgagggtgag ttcctggtgt tggatcagcg ggcagccgac 540
tacaaccagg ccctgggcac ctgtcgcctg gcaggcacag cgctctgtgt ggcagctgga 600
gttctgctcg ccatctgcct cttctgggcc atgataggct ggctgagcca ggacaccaag 660
gcagagccct tggaccccga agccgacagc cacgtggagg tcttcgggga tgagccagag 720
cagcagttgt cacccatttt ccgcaatgcc agtggccagt catggttctc gccacccgcc 780
agcccctttg ggcaatcttc tgtgcagact atccagccca agagggactc ctgagctgcc 840
cacatggcct aagatgtggg tcctggatcc ttccccttc tcaccataac cccctctcag 900
tgtttcccca acttctccct tttagcaggg tccctttaga gcccaactcc aggtcaaatc 960
tggagctcaa atcccagtgc tccctcccca ggagtggggc cccaactctt ccaagatacc 1020
agcattcctc aagtcctccc aaaacttcct acccacaccc tcttcccaag gccctcaggg 1080
gcagaaaaca tctccttcaa cccgtcccca ctccttcctc tgcatgacct tgggcaaacc 1140
cttgcccttt caagccatca gctcctgcct ctctgccatg agggctttgg atcagattcc 1200
tcttctcgcc aggatgagga cacgcactgc cctccataga cacagatgaa ggggtggggg 1260
tcattcagct cgaatgggtc ccagatgctc acttggcctt tccctgcagg atgagtgaag 1320
acgtttgcct ctcacagtgt gtcttctacc tgcattttgg catcagagcc ccccagccca 1380
cccaccacag gcaattacta gccctagttg ataggtgagg tgggtgaaga aggctggagg 1440
tgacatgtcc gaggtcacac aacaaagcag catgcaggaa ctagaaacac atcttcagcc 1500
tcctcctggg ccagctcttg tgctacaggt ggggcggagc cagcccctca ccttcctggt 1560
tccctgaggg tcctcagggt ggaggacagg tttggcccag aaagactagc cagaggcctg 1620
atggtcccag gtggctctgg atatactttg gatatggatt taaatggtct ctaagagccg 1680
ggggtagggg gcaggaaaag tgggttgtct ttgcccctca aagtccacct acctagaaac 1740
caagcccacg gtcttggccg tgaccctgat aataaatggg ctctctcaga ggcgccagcc 1800
cctccctccc cagccggagg cgtcatctct cttctgtacc actagaggga gctctgatgc 1860
agctggagag cagcgctcaa ggctctcgcc cctcccctcc ctaaccctta ccttcagtct 1920
ccaccagcct gaagggcctc ctaggggatc ctcaggcggc ccccaccagg gcacacccta 1980
ctgtccttgt gcctcacgcc ccctcctcat cctgcacccc ttccatccca ccttcccttt 2040
caataaacag ctgggatgga tactgaaaaa aaaaaaaaaa 2080

<210> 51
<211> 1420
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3729267CB1

<400> 51
cttccagcgc cgcttgcgct ccggagcgct ggctctgctg gcgctgagaa atggaccaat 60
tttgacaaga tatagtgctg cagcgtgcct gatgggatat attcagtcat ggcgtccgaa 120
ctttgtaaga cgatctctgt ggcaaggcta gaaaagcaca agaatttgtt cttaaattat 180

aggaatctgc accattttcc attggagtta ctgaaagatg agggactgca gtacttggag 240
agactctata tgaaaaggaa ctccctgaca tccttgccag aaaaccttgc tcagaagctt 300
ccaaaccttg tggaactata cctgcactca aataacatag ttgtggttcc ggaagccatt 360
gggtctcttg taaaactcca atgtctggat cttagtgaca atgccttaga aattgtttgc 420
ccagaaattg gtcgtctgag agctttacgt catcttcgat tagctaataa ccaactgcaa 480
ttcctacctc cagaggttgg cgatttgaag gagctgcaga cactagacat ttctaccaat 540
cgtttgctaa ctttacccga gaggcttcac atgtgccttt ctctgcagta cctcactgtg 600
gaccgaaatc gtctatggta tgtgccgcgc catctctgcc agctgcccag cctcaatgag 660
ctctccatgg ctggaaaccg tcttgcattt ttgccacttg atttaggtcg atctcgagaa 720
ctacagtatg tatacgtgga taacaacatt cacctgaaag gcttgccatc ttatctgtac 780
aataaagtca tcgggtgcag tggctgtggt gctcccattc aagtttccga ggtgaagctg 840
ctttcctttt catcagggca gcgaaccgtt ttcctcccag ctgaggtgaa ggccataggg 900
acggagcatg atcacgtcct ccctctgcag gaattggcta tgagagggct gtatcatacc 960
taccacagct tgctgaaaga tttgaacttt ctgtctccaa tctcattacc cagaagtctc 1020
ctagagctgc tgcactgccc tctggggcac tgtcatcggt gtagtgagcc tatgtttacc 1080
atcgtctacc ccaagctctt tcccttgaga gagacgccaa tggcagggct gcaccagtgg 1140
aagacaactg ttagttttgt ggcttactgc tgctccaccc agtgtctgca gacttttgac 1200
ctgctgagtt gataaacact caagaacctc aggagcgctg ccagcttgac actggggaat 1260
ccagccagtc cagcacactc ttccatcctg tcctgtccaa tgcgggggca ctgcagaact 1320
ctctagaaat gtcatgattg agcttcagag ctaaaatgcc ttcacccttc ccccaagttg 1380
gaatatatcc tcccccaaat taaggaaaaa aaaaaaaaaa 1420

<210> 52
<211> 2703
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3768771CB1

<400> 52
gctgattacc ttcctttaaa ccagtccatc taaaggttgg tgttaagagg cttcccatgg 60
gaggggtggc cccaggtaat acccatcttc ctctccccag gctctgaaga gcggccattc 120
ctcagattcg aagctgaaca catctccaac tacacagccc ttctgctgag cagggatggc 180
aggaccctgt acgtgggtgc tcgagaggcc ctctttgcac tcagtagcaa cctcagcttc 240
ctgccaggcg gggagtacca ggagctgctt tggggtgcag acgcagagaa gaaacagcag 300
tgcagcttca agggcaagga cccacagcgc gactgtcaaa actacatcaa gatcctcctg 360
ccgctcagcg cagntcacct gttcacctgt ggcacagcag ccttcagccc catgtgtacc 420
tacatcaaca tggagaactt caccctggca agggacgaga aggggaatgt cctcctggaa 480
gatggcaagg gccgttgtcc cttcgacccg aatttcaagt ccactgccct ggtggttgat 540
ggcgagctct acactggaac agtcagcagc ttccaaggga atgacccggc catctcgcgg 600
agccaaagcc ttcgccccac caagaccgag agctccctca actggctgca agacccagct 660
tttgtggcct cagcctacat tcctgagagc ctgggcagct tgcaaggcga tgatgacaag 720
atctactttt tcttcagcga gactggccag gaatttgagt tctttgagaa caccattgtg 780
tcccgcattg cccgcatctg caagggcgat gagggtggag agcgggtgct acagcagcgc 840
tggacctcct tcctcaaggc ccagctgctg tgctcacggc ccgacgatgg cttccccttc 900
aacgtgctgc aggatgtctt cacgctgagc cccagccccc aggactggcg tgacaccctt 960
ttctatgggg tcttcacttc ccagtggcac aggggaacta cagaaggctc tgccgtctgt 1020
gtcttcacaa tgaaggatgt gcagagagtc ttcagcggcc tctacaagga ggtgaaccgt 1080
gagacacagc agtggtacac cgtgacccac ccggtgccca caccccggcc tggagcgtgc 1140
atcaccaaca gtgcccggga aaggaagatc aactcatccc tgcagctccc agaccgcgtg 1200
ctgaactttc tcaaggacca cttcctgatg gacgggcagg tccgaagccg catgctgctg 1260
ctgcagcccc aggctcgcta ccagcgcgtg gctgtacacc gcgtccctgg cctgcaccac 1320

acctacgatg tcctcttcct gggcactggt gacggccggc tccacaaggc agtgagcgtg 1380
ggcccccggg tgcacatcat tgaggagctg cagatcttct catcgggaca gcccgtgcag 1440
aatctgctcc tggacaccca caggggctg ctgtatgcgg cctcacactc gggcgtagtc 1500
caggtgccca tggccaactg cagcctgtac cggagctgtg gggactgcct cctcgcccgg 1560
gacccctact gtgcttggag cggctccagc tgcaagcacg tcagcctcta ccagcctcag 1620
ctggccacca ggccgtggat ccaggacatc gagggagcca gcgccaagga cctttgcagc 1680
gcgtcttcgg ttgtgtcccc gtcttttgta ccaacagggg agaagccatg tgagcaagtc 1740
cagttccagc ccaacacagt gaacactttg gcctgcccgc tcctctccaa cctggcgacc 1800
cgactctggc tacgcaacgg ggcccccgtc aatgcctcgg cctcctgcca cgtgctaccc 1860
actggggacc tgctgctggt gggcacccaa cagctggggg agttccagtg ctggtcacta 1920
gaggagggct tccagcagct ggtagccagc tactgcccag aggtggtgga ggacggggtg 1980
gcagaccaaa cagatgaggg tggcagtgta cccgtcatta tcagcacatc gcgtgtgagt 2040
gcaccagctg gtggcaaggc cagctggggt gcagacaggt cctactggaa ggagttcctg 2100
gtgatgtgca cgctctttgt gctggccgtg ctgctcccag ttttattctt gctctaccgg 2160
caccggaaca gcatgaaagt cttcctgaag caggggggaat gtgccagcgt gcaccccaag 2220
acctgccctg tggtgctgcc ccctgagacc cgcccactca acggcctagg gcccccctagc 2280
accccgctcg atcaccgagg gtaccagtcc ctgtcagaca gcccccccggg gtcccgagtc 2340
ttcactgagt cagagaagag gccactcagc atccaagaca gcttcgtgga ggtatcccca 2400
gtgtgccccc ggccccgggt ccgccttggc tcggagatcc gtgactctgt ggtgtgagag 2460
ctgacttcca gaggacgctg ccctggcttc aggggctgtg aatgctcgga gagggtcaac 2520
tggacctccc ctccgctctg ctcttcgtgg aacacgaccg tggtgcccgg cccttgggag 2580
ccttggagcc agctggcctg ctgctctcca gtcaagtagc gaactcctac cacccagaca 2640
cccaaacagc cgtggcccca gaggtcctgg ccaaatatgg gggcctgcct aggttggtgg 2700
aaa 2703

<210> 53
<211> 571
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 4248993CB1

<400> 53
ctgtgttgag taaccatggg gaggaagctg gacctgtctg gtttgactga tgatgaaaca 60
gagcatgttc ttcaggtggt tcaaagagac ttcaatcttc gcaaaaaaga agaagaacga 120
ctaagtgagc tgaagcagaa gctggatgag gaaggcagca agtgcagcat cctctcgaag 180
caccagcagt ttgtggagca ctgctgcatg cgctgctgct cgcccttcac cttcctcgtc 240
aacaccaagc gccagtgtgg agattgcaaa ttcaatgtct gcaagagctg ctgctcctac 300
cagaagcacg aaaaggcctg ggtctgctgc gtctgccagc aagcgaggct tctgagggcc 360
caatctctgg aatggttcta caataatgtg aagagccgct tcaagcgctt tggcagtgca 420
aggttctgaa gaacctgtac agggagcacc ggctggagag tggcgcgtgc ttcgacattc 480
taggaggaag ctttttgacg tcaaccctgg agatgaggga gcattctggc agtgattcac 540
atttataggc agtcagagga catagtgtgt g 571

<210> 54
<211> 1293
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 5402418CB1

<400> 54
cccagctcta gcgaaaagcc gccggtattt ctccatctgg ctctcctcta cctccaggca 60
ggctcacccg agatccccgc cccgaacccc ccctgcacac tcggcccagc gctgttgccc 120
ccggagcgga cgtttctgca gctattctga gcacaccttg acgtcggctg agggagcggg 180
acagggtcag cggcgaagga ggcaggcccc gcgcggggat ctcggaagcc ctgcggtgca 240
tcatgaagtt ccagtacaag gaggaccatc cctttgagta tcggaaaaag gaaggagaaa 300
agatccggaa gaaatatccg gacagggtcc ccgtgattgt agagaaggct ccaaaagcca 360
gggtgcctga tctggacaag aggaagtacc tagtgccctc tgaccttact gttggccagt 420
tctacttctt aatccggaag agaatccacc tgagacctga ggacgcctta ttcttctttg 480
tcaacaacac catccctccc accagtgcta ccatgggcca actgtatgag gacaatcatg 540
aggaagacta ttttctgtat gtggcctaca gtgatgagag tgtctatggg aaatgagtgg 600
ttggaagccc agcagatggg agcacctgga cttgggggta ggggaggggt gtgtgtgcgc 660
gacatgggga aagagggtgg ctcccaccgc aaggagacag aaggtgaaga catctagaaa 720
cattacacca cacacaccgt catcacattt tcacatgctc aattgatatt ttttgctgct 780
tcctcggccc agggagaaag catgtcagga cagagctgtt ggattggctt tgatagagga 840
atggggatga tgtaagttta cagtattcct ggggtttaat tgttgtgcag tttcatagat 900
gggtcaggag gtggacaagt tggggccaga gatgatggca gtccagcagc aactccctgt 960
gctcccttct ctttgggcag agattctatt tttgacattt gcacaagaca ggtagggaaa 1020
ggggacttgt ggtagtggac catacctggg gaccaaaaga gacccactgt aattgatgca 1080
ttgtggcccc tgatcttccc tgtctcacac ttcttttctc ccatcccggt tgcaatctca 1140
ctcagacatc acagtaccac cccaggggtg gcagtagaca acaacccaga aatttagaca 1200
gggatctctt acctttggaa aataggggtt aggcatgaag gtggttgtga ttaagaagat 1260
ggttttgtta ttaaataaca ttaaactagg att 1293

<210> 55
<211> 375
<212> PRT
<213> Homo sapiens

<300>
<308> g3002527

<400> 55
Met Glu Phe Ser Leu Leu Leu Pro Arg Leu Glu Cys Asn Gly Ala
1 5 10 15
Ile Ser Ala His Arg Asn Leu Arg Leu Pro Gly Ser Ser Asp Ser
20 25 30
Pro Ala Ser Ala Ser Pro Val Ala Gly Ile Thr Gly Met Cys Thr
35 40 45
His Ala Arg Leu Ile Leu Tyr Phe Phe Leu Val Glu Met Glu Phe
50 55 60
Leu His Val Gly Gln Ala Gly Leu Glu Leu Pro Thr Ser Asp Asp
65 70 75
Pro Ser Val Ser Ala Ser Gln Ser Ala Arg Tyr Arg Thr Gly His
80 85 90
His Ala Arg Leu Cys Leu Ala Asn Phe Cys Gly Arg Asn Arg Val
95 100 105
Ser Leu Met Cys Pro Ser Trp Ser Pro Glu Leu Lys Gln Ser Thr
110 115 120
Cys Leu Ser Leu Pro Lys Cys Trp Asp Tyr Arg Arg Ala Ala Val

125 130 135
Pro Gly Leu Phe Ile Leu Phe Phe Leu Arg His Arg Cys Pro Thr
140 145 150
Leu Thr Gln Asp Glu Val Gln Trp Cys Asp His Ser Ser Leu Gln
155 160 165
Pro Ser Thr Pro Glu Ile Lys His Pro Pro Ala Ser Ala Ser Gln
170 175 180
Val Ala Gly Thr Lys Asp Met His His Tyr Thr Trp Leu Ile Phe
185 190 195
Ile Phe Ile Phe Asn Phe Leu Arg Gln Ser Leu Asn Ser Val Thr
200 205 210
Gln Ala Gly Val Gln Trp Arg Asn Leu Gly Ser Leu Gln Pro Leu
215 220 225
Pro Pro Gly Phe Lys Leu Phe Ser Cys Pro Ser Leu Leu Ser Ser
230 235 240
Trp Asp Tyr Arg Arg Pro Pro Arg Leu Ala Asn Phe Phe Val Phe
245 250 255
Leu Val Glu Met Gly Phe Thr Met Phe Ala Arg Leu Ile Leu Ile
260 265 270
Ser Gly Pro Cys Asp Leu Pro Ala Ser Ala Ser Gln Ser Ala Gly
275 280 285
Ile Thr Gly Val Ser His His Ala Arg Leu Ile Phe Asn Phe Cys
290 295 300
Leu Phe Glu Met Glu Ser His Ser Val Thr Gln Ala Gly Val Gln
305 310 315
Trp Pro Asn Leu Gly Ser Leu Gln Pro Leu Pro Pro Gly Leu Lys
320 325 330
Arg Phe Ser Cys Leu Ser Leu Pro Ser Ser Trp Asp Tyr Gly His
335 340 345
Leu Pro Pro His Pro Ala Asn Phe Cys Ile Phe Ile Arg Gly Gly
350 355 360
Val Ser Pro Tyr Leu Ser Gly Trp Ser Gln Thr Pro Asp Leu Arg
365 370 375

<210> 56
<211> 309
<212> PRT
<213> Homo sapiens

<300>
<308> g847722

<400> 56
Met Val Ser Phe Val Ser Asn Tyr Ser His Thr Ala Asn Ile Leu
1 5 10 15
Pro Asp Ile Glu Asn Glu Asp Phe Ile Lys Asp Cys Val Arg Ile
20 25 30
His Asn Lys Phe Arg Ser Glu Val Lys Pro Thr Ala Ser Asp Met
35 40 45
Leu Tyr Met Thr Trp Asp Pro Ala Leu Ala Gln Ile Ala Lys Ala
50 55 60
Trp Ala Ser Asn Cys Gln Phe Ser His Asn Thr Arg Leu Lys Pro
65 70 75
Pro His Lys Leu His Pro Asn Phe Thr Ser Leu Gly Glu Asn Ile

```
                80                  85                  90
Trp Thr Gly Ser Val Pro Ile Phe Ser Val Ser Ser Ala Ile Thr
                95                  100                 105
Asn Trp Tyr Asp Glu Ile Gln Asp Tyr Asn Phe Lys Thr Arg Ile
                110                 115                 120
Cys Lys Lys Val Cys Gly His Tyr Thr Gln Val Val Trp Ala Asp
                125                 130                 135
Ser Tyr Lys Val Gly Cys Ala Val Gln Phe Cys Pro Lys Val Ser
                140                 145                 150
Gly Phe Asp Ala Leu Ser Asn Gly Ala His Phe Ile Cys Asn Tyr
                155                 160                 165
Gly Pro Gly Gly Asn Tyr Pro Thr Trp Pro Tyr Lys Arg Gly Ala
                170                 175                 180
Thr Cys Ser Ala Cys Pro Asn Asn Asp Lys Cys Leu Asp Asn Leu
                185                 190                 195
Cys Val Asn Asp Ser Glu Thr Lys Ser Asn Val Thr Thr Met Leu
                200                 205                 210
Tyr Ile Arg Leu Ala His Ile Ser Thr
                215
```